



(19) Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 1 035 207 A1

36  
2  
P

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
13.09.2000 Bulletin 2000/37

(21) Application number: 99104664.0

(22) Date of filing: 09.03.1999

(51) Int Cl.7: C12N 15/53, C12N 15/11,  
C12N 15/85, C12N 9/02,  
C12N 5/10, C12Q 1/02,  
C07K 16/40, A61K 39/395,  
A61K 38/44, A01K 67/027,  
G01N 33/50, G01N 33/53

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE  
Designated Extension States:  
AL LT LV MK RO SI

(71) Applicant: MultiGene Biotech GmbH  
97074 Würzburg (DE)

(72) Inventors:  
• Weber, Bernhard H.F.  
97218 Gerbrunn (DE)

• Marquardt, Andreas  
97218 Gerbrunn (DE)

(74) Representative: Schmidt, Werner, Dr.  
Robert-Bunsen-Strasse 15  
65929 Frankfurt am Main (DE)

### Remarks:

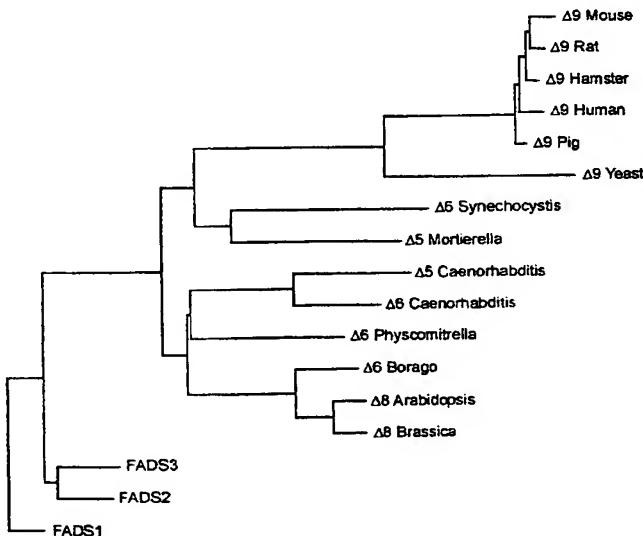
The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) cDNA molecules of the members of gene family encoding human fatty acid desaturases and their use in diagnosis and therapy

(57) The present invention relates to the cloning and sequencing of the cDNA molecules of three members of a gene family encoding three human fatty acid desaturases, fatty acid desaturase-1 (FADS 1), fatty acid desaturase-2 (FADS2) and fatty acid desaturase-3 (FADS3). The invention also relates to diagnostic meth-

ods of screening for and detection of FADS1, FADS2, FADS3 and gene therapy utilizing recombinant DNA as well as the generation of animal models (knock-in, knock-out, transgenic animals), anti-FADS1, anti-FADS2, anti-FADS3 antibodies and use in screenings for modulating drugs.

Fig.2



**Description****Field of the invention**

5 [0001] The present invention relates to the cloning and sequencing of the cDNA molecules of three members of a gene family encoding three human fatty acid desaturases, fatty acid desaturase-1 (FADS1), fatty acid desaturase-2 (FADS2) and fatty acid desaturase-3 (FADS3). The invention also relates to diagnostic methods of screening for and detection of FADS1, FADS2, FADS3 and gene therapy utilizing recombinant DNA as well as the generation of animal models (knock-in, knock-out, transgenic animals), anti-FADS1, anti-FADS2, anti-FADS3 antibodies and use in screenings for modulating drugs.

10 [0002] Cellular membranes are dynamic structures in which variable amounts of proteins are embedded in a lipid bilayer whose hydrophobic characteristics are largely due to fatty acid moieties of complex lipids (Singer and Nicolson 1972). The 'fluidity' of the membranes are achieved by incorporating unsaturated fatty acyl chains of varying lengths and varying degrees of unsaturation into the lipids (Stubbs and Smith 1984). In animals, some of the unsaturated fatty acids need to be supplied by the diet ('essential polyunsaturated fatty acids') but, in part, can also be synthesized de novo by oxidative desaturation (i.e. formation of double bonds) of saturated fatty acids of plant and animal origin. 15 Polyunsaturated fatty acid formation requires acetyl-CoA dependent chain elongation and desaturation. Most mammalian tissues can modify acyl chains by introducing more than one double bond with the first one generally at the Δ-9 position between carbons C-9 and C-10. Subsequent double bonds may then be inserted at the Δ-4, Δ-5, and Δ-6 positions by individual desaturase activities (Cook 1991).

20 [0003] For the two major precursors of the (n-6) and (n-3) series of polyunsaturated fatty acids, linoleic 18:2(n-6) and alpha-linolenic 18:3(n-3) acids, animals depend entirely on their dietary intake. By alternating sequences of desaturation (involving the subsequent action of Δ4, Δ5- and Δ6-desaturases, respectively) and C2 chain elongation, linoleic and alpha-linolenic acids are utilized to form arachidonic acid, 20:4(n-6), and the (n-3) acyl chains eicosapentaenoic acid, 20:5(n-3), and docosahexaenoic acid, 22:6(n-3), respectively (Cook 1991).

25 [0004] Linoleic and arachidonic acid are the only members of the (n-6) family that accumulate in large quantities in liver and most other animal tissues. The intermediates 18:3(n-6) and 20:3(n-6) are formed from 18:2(n-6) by Δ6-desaturation, chain elongation and Δ5-desaturation (Horrobin 1993). As a component of phospholipids arachidonic acid is abundant in cellular membranes but also serves as the primary precursor of oxygenated derivatives such as prostaglandine E2 which is pro-inflammatory and regulates cell function of the immune system.

30 [0005] The (n-3) acyl chains eicosapentaenoic acid [20:5(n-3)] and docosahexaenoic acid [22:6(n-3)] are most abundant in cerebral cortex, retina, and spermatozoa. Although it is generally assumed that the liver is the major source of 22:6(n-3), it has been shown that docosahexaenoic acid can also be produced by retinal pigment epithelium (Wang and Anderson 1993) as well as brain astrocytes (Moore et al. 1991, Delton-Vandenbrouke et al. 1997). In retinal rod outer segments, phospholipids may contain 40-60% of 22:6(n-3) which can markedly influence membrane fluidity due to the presence of six double bonds.

35 [0006] In recent years there has been increasing interest in the role of polyunsaturated fatty acids in the pathobiology of a number of chronic conditions such as coronary and peripheral vascular disease (Horrobin 1995), acute and chronic inflammatory immune responses (Calder 1998, Fan and Chapkin 1998, Grimble and Tappia 1998), cutaneous abnormalities (Horrobin 1989, Grattan et al. 1990), essential hypertension (Russo et al. 1997, Chi and Gupta 1998), diabetes mellitus (Mori et al. 1997), asthma (Leichsenring et al. 1995, Villani et al. 1998, Hodge et al. 1998) and rheumatoid 40 arthritis (James and Cleland 1997, Ariza-Ariza et al. 1998, Grimble and Tappia 1998). A particular role has been attributed to gamma-linolenic acid [18:3(n-6)] as an anti-cancer polyunsaturated fatty acid. It has been shown that 18:3 (n-6) confers anticancer properties by a variety of mechanisms such as (i) up-regulation of E-cadherin, a cell-cell adhesion molecule which acts as a suppressor of metastasis (Jiang et al. 1995), (ii) regulation of desmosome-mediated cell-cell adhesion in human cancer cells (Jiang et al. 1997a), (iii) up-regulation of the metastasis-suppressor gene nm-23 thus contributing to the inhibition of the in vitro invasion of tumor cells (Jiang et al. 1998a), (iv) up-regulation of maspin expression, a mammary serine protease inhibitor, with profound effects on motility of cancer cells (Jiang et al. 1997b) and (v) finally inhibition of cell cycle progression via regulation of phosphorylation and subsequent degradation of cell cycle inhibitors p27kip1 and p57kip2 (Jiang et al. 1998b).

45 [0007] To further understand lipid-related function in human health and disease additional research into fatty acid biosynthesis and metabolism is required. In particular, we need to understand the pharmacological properties, the mechanisms of action and the tissue-specific regulation of composition of the polyunsaturated fatty acids and their metabolites. This will provide additional insight into the role of the polyunsaturated fatty acids in various chronic disease states and will make it feasible to focus pharmacogenomic research on drug design and valuation with the goal of

ameliorating acute health problems associated with impaired lipid function. As a prerequisite, the genes and their gene products involved in the above-mentioned processes need to be identified and characterized.

[0008] It is the objective of the present invention to provide cDNA molecules of three novel members of the human membrane fatty acid desaturase gene family, termed FADS1, FADS2 and FADS3. The three genes share a nucleic acid identity of approximately 50-60% and an amino acid identity of about 77% with each other. Similar to other membrane-bound desaturases from mammals, fungi, insects, plants and cyanobacteria FADS1, FADS2 and FADS3 reveal a hydropathy profile typical of membrane-bound desaturases and share three regions of highly conserved primary sequence of the general histidine motif HX<sub>2(3)</sub>H (Shanklin et al. 1994). The histidine residues may act as metal-chelating ligands involved in the binding of oxygen in the reaction center (Shanklin et al. 1995). Together, these features confirm FADS1, FADS2 and FADS3 as novel members of the desaturase family of fatty acyl chain-modifying enzymes.

[0009] Amino acid identity of FADS1, FADS2 and FADS3 to known desaturases (e.g. from *Arabidopsis thaliana*, *Brassica napus*, *Synechocystis spec.*, *Borago officinalis*, *Helianthus annuus*, *Saccharomyces cerevisiae* and *Caenorhabditis elegans*) is restricted to the respective carboxy terminal regions (amino acid positions 260 to 422) revealing an overall sequence identity of approximately 27%. Interestingly, the respective amino-termini of the three novel proteins demonstrate similarities to cytochrome b5 (amino acid positions 4 to 75; Fig. 1). Cytochrome b5 is a small hemoprotein and functions as an intermediate donor in a number of oxidation/reduction reactions including e.g. the NADH-dependent Δ9 stearoyl-CoA desaturation (Strittmatter et al. 1974) or the Δ5 desaturation in cholesterol biosynthesis (Reddy et al. 1977). From the amino acid alignments we conclude that FADS1, FADS2 and FADS3 are fusion proteins consisting of a N-terminal cytochrome b5 and a C-terminal desaturase-like enzyme. From a functional point of view, this fusion of two activities may increase the efficiency of electron transport required for desaturation by covalently bringing together the presumed electron donor (cytochrome b5) and its putative acceptor (desaturase-like enzyme). Other heme fusion proteins containing the cytochrome b5 domain have been identified and represent a superfamily of fused proteins (Guillard and Lederer 1979). Besides others this superfamily includes the yeast flavocytochrome b<sub>2</sub>, sulfite oxidase, nitrate reductase, the yeast Δ9 acyl-CoA desaturase and more recently the sunflower cytochrome b5-desaturase fusion protein (Sperling et al. 1995). The three novel desaturase-like enzymes reported herein, FADS1, FADS2 and FADS3, can be added to the growing list of members of this superfamily of fused proteins (Fig. 2).

#### Summary of the invention

[0010] The eukaryotic fatty acid desaturases represent a group of iron-containing enzymes that catalyze NAD(P)H- and O<sub>2</sub>-dependent introduction of double bonds into fatty acyl chains. Impairment of desaturase activities has been implicated in a variety of human conditions including liver disease, coronary artery disease and cancer. With the present invention we are providing three isolated human cDNA molecules that encode three novel members of a cytochrome-b5-containing fusion protein with similarity to plant and lower animal desaturase enzymes, termed fatty acid desaturase-1 (FADS1) (represented by Fig. 3 and SEQ ID NO. 1), fatty acid desaturase-2 (FADS2) (represented by Fig. 4 and SEQ ID NO. 2) and fatty acid desaturase-3 (FADS3) (represented by Fig. 5 and SEQ ID NO. 3).

#### FADS1 protein

[0011] MAPDPVAAETAACQGPTPRYFTWDEVAQRSGCEERWLVIDRKVYNISEFTRRHPGGS RVISHYAGQDATDP-FVAFHINKGLVKYMNSSLIGELSPEQPSFEPTKNKELTDEFREL RATVERMGLMKANHVFFLLYLHILLLDGAALWTL-WVFGTSFLPFLCAVLLSAVQAQA GWLQHDFGHLSVFSTSKWNHLLHHFVIGHLKgapaswwnNMHFQHHAKPNC-FRKD PDINMHPPFFFALGKILSVELGKQKKKYMYPNHQHKYFFLIGPPALLPLYFQWYIFYFVIQ RKKWVDLAWMITFY-VRFFLTYVPLLGLKAFLGLFFIVRFLESNWFVWVTQMNHIPMHID HDRNMDWVSTQLQATCNVHKSAFNDWFSGL-NFQIEHHLFPTMPRHNYHKVAPLVQ SLCAKHGIEYQSKPLLSAFADIIHSLKESGQLWLDAYLHQ

#### FADS2 protein

[0012] MGKGGNQGEGAAEREVSPTFSWEIQKHNLRTDRWLVIDRKVYNITKWSIQHPGG QRVIGHYAGEDAT-DAFRAFHPDLEFGKFKLPLLIGEAAPEEPSQDHGKNSKITEDFRA LRKTAEDMNLFKTNHVFFLLLAHIalesia-WFTVFYFGNGWIPTLITAFVLATSOAQAG WLQHDYGHLSWRKPKWNHLVHKFVIGHLKgasanwwnHRHFQHH-AKPNIFHKDPD VNMLHVFVLGEWQPIEYGKKKLKYLPYNHQHEYFFLIGPPLLIPMYFQYQIIMTMIVHKN WVDLA-WAVSYYIRFFITYIPFYGILGALLFLNFIRFLESHWFVdNTQMNHIVMEIDQEAY RDWFSSQLTATCNVEQSFFNDWFS-GHlnfQIEHHLFPTMPRHNLHKIAPLVKSLCAK HGIEYQEPLLRAALDIIRSLKSGKLWLDAYLHK

## FADS3 protein

[0013] MGGVGEPPGPREGPAQPGAPLPTFCWEQIRAHQDQPGDKWLVIERRVYDISRWAQRHP GGSRLIGHHGAEDATDAFRFHQDLNFKFLQPLLIGELAPEEPSQDGPLNAQLVED FRALHQAAEDMKLFDSAFTFAFLLGHLAMEVLAWLLIYLLGPGWVPSALAAFLAISQ AQSWCLQHDLGHASIFKKSSWWNHVAQKFMGQLKGFSAHWWNFRHFQHAKPNIF HKDPDVTVAPVFLGESSVEYGKKRRLPYNQQHLYFFLIGPPLTLVNFEVENLAY MLVCMQWADLLWAASFYARFFLSYLPFYGVPGVLLFFVAVRVLESHWFVWITQMNHI PKEIGHEKHRDWVSSQLAATCNVEPSLTNWFGHHLNFQIEHHLFPRMPRHNSRVA PLVKSLCAKHGLSYEVKPFLTALVDIVRSLKSGDIWLDAYLHQ

[0014] Studies to clarify the specificity and the subcellular location of these ubiquitously expressed fusion proteins are in progress. Also, the detailed cellular functions and dysfunctions of the desaturase-like domains are being investigated in appropriate cellular and animal systems. This will address the question whether and to which extent these novel enzymes are involved in human disease. The invention encompasses the three cDNA molecules, FADS1, FADS2, and FADS3, the nucleotide sequence of these cDNAs, and the putative amino acid sequences of the FADS1 (represented by Fig. 6 and SEQ ID NO. 4), FADS2 (represented by Fig. 7 and SEQ ID NO. 5), and FADS3 represented by Fig. 8 and SEQ ID NO. 6) proteins.

[0015] Also comprehended by this invention are oligonucleotide primers comprising the cDNA molecule or its complementary strand allowing the amplification of FADS1 (represented by Fig. 9 and SEQ ID NOS. 7-12), FADS2 (represented by Fig. 9 and SEQ ID NOS. 13-18), and FADS3 (represented by Fig. 9 and SEQ ID NOS. 19-22), by the reverse transcriptase polymerase chain reaction (RT-PCR). Such primers are particularly useful and will provide researchers and physicians with an enhanced ability to assess the role of FADS1, FADS2, and FADS3 in human disease. The present invention also relates to methods of screening for and detection of FADS1, FADS2, and FADS3 mutation carriers including prenatal FADS1, FADS2, and FADS3 screening and diagnosis.

[0016] Having provided the isolated human FADS1, FADS2, and FADS3 cDNA sequences, also comprehended by this invention are the FADS1, FADS2, and FADS3 proteins, and derivatives thereof, in aspects of diagnosis and treatment of human disease. Finally, the invention pertains to proteins which comprise the same or substantially the same amino acid sequence (at least 200 amino acids) as that represented by Figs. 6, 7, 8 and SEQ ID NOS. 4, 5, 6 or a variant of the amino acid sequences having a deletion, addition or substitution of 1 to 10 amino acids, or its salt.

[0017] Another aspect of the invention is the use of the FADS1, FADS2, and FADS3 proteins as a target for drug and gene therapy in the treatment of human disease. This includes the generation and utilization of FADS1, FADS2, and FADS3-targeted animal models (knock-in, knock-out, transgenic animals) and anti-FADS1, -FADS2, and -FADS3 antibodies that specifically detect the FADS1, FADS2, and FADS3 proteins, respectively.

[0018] The foregoing and other features and advantages of the invention will become more apparent from the following detailed description and accompanying drawings.

[0019] One aspect of the invention are the isolated cDNAs selected from the group consisting of:

- (a) a polynucleotide having at least a 65 % homology, preferably at least a 80 % homology with a polynucleotide encoding a polypeptide selected from the group consisting of the polypeptides of SEQ ID NO: 4, SEQ ID NO: 5 and SEQ ID NO: 6;
- (b) a polynucleotide having at least a 65 % homology, preferably at least a 80 % homology with a polynucleotide which by virtue of the redundancy of the genetic code, encodes the same polypeptide selected from the group consisting of the polypeptides of SEQ ID NO: 4, SEQ ID NO: 5 and SEQ ID NO: 6;
- (c) a DNA molecule capable of hybridization under stringent conditions to a DNA molecule according to (a) or (b);
- (d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and
- (e) a oligonucleotide comprising at least 15 consecutive nucleotides of the polynucleotide of (a), (b), (c) or (d)

(including DNAs which are synonymous to the DNAs of (a), (b), (c), (d) and (e) due to the degeneracy of the genetic code)

especially isolated cDNAs selected from the group consisting of:

- (a) a polynucleotide having at least a 65 % homology, preferably at least a 80 % homology with a polynucleotide sequence selected from the group consisting of the polynucleotides of SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 3;
- (b) a DNA molecule capable of hybridization under stringent conditions to a DNA molecule according to (a);
- (c) a polynucleotide which is complementary to the polynucleotide of (a) or (b);
- (d) a oligonucleotide comprising at least 15 consecutive nucleotides of the polynucleotide of (a), (b) or (c); and
- (e) a DNA which is synonymous to the DNAs of (a), (b), (c) or (d) due to the degeneracy of the genetic code.

[0020] In the scope of the invention are polynucleotides having a polynucleotide encoding a polypeptide selected

from the group consisting of the polypeptides of SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 3 and polynucleotides having a polynucleotide sequence selected from the group consisting of the polynucleotides of SEQ ID NO: 4, SEQ ID NO: 5 and SEQ ID NO: 6, but DNAs comprising a nucleotide sequence with at least a 65 % homology with these nucleotide sequences is also within the scope of the invention.

[0021] Furthermore within the scope of the invention are:

[0022] A recombinant vector comprising the disclosed DNA molecules.

[0023] Transgenic host cells such as COS7, fibroblast cell lines or any other tissue-specific cell lines, as well as a transgenic host cell transformed by the DNA or the vector, a corresponding transgenic organism or a corresponding transgenic knock-in or knock-out animal model.

[0024] Polypeptides and corresponding proteins comprising at least 65 %, preferably 85 %, especially 100 % of a polypeptide sequence selected from the group consisting of the polypeptides of SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 3; polypeptides comprising a polypeptide sequence with at least a 65 % homology with the said polypeptides; peptides comprising at least 15, preferably 30, especially 60 consecutive amino acids of the said polypeptides; and polypeptides having substantially the same amino acid sequence as the said polypeptides, or having a variant of the amino acid sequence of the polypeptides with a deletion, addition or substitution of 1 to 10 amino acids. The salts of the peptides and proteins are also within the scope of the invention.

[0025] A process for preparing the proteins which comprises cultivating the transformants to form the proteins.

[0026] A method of screening for modulators in well known assays using constructs such as FADS1, FADS2, and FADS3 promoter luciferase or green fluorescent protein hybrids or screening for interacting proteins or factors using state of the art technologies like the interaction trap technology to screen for interacting substances of FADS1, FADS2, and FADS3 or isolated domains of FADS1, FADS2, and FADS3.

#### A method of screening chemical libraries comprising transformed cell lines

[0027] A compound which alters 1 reacts with at least one epitope of the proteins and which is obtained by screening methods utilizing the FADS1, FADS2, and FADS3 cDNAs or protein molecules.

[0028] Use of antibodies against the FADS1, FADS2, and FADS3 proteins for diagnostic or therapeutic purposes.

[0029] A pharmaceutical composition comprising as an effective component of the proteins or a partial peptide of the proteins, and a pharmaceutically acceptable carrier or diluent.

[0030] The term "knock-out animal" as used herein is intended to describe an animal containing a gene which has been modified by homologous recombination. The homologous recombination event may completely disrupt the gene such that a functional gene product can no longer be produced (hence the name "knock-out") or the homologous recombination event may modify the gene such that an altered, although still functional, gene product is produced.

[0031] The term "knock-in" as used herein is intended to describe a variation of gene targeting that uses homologous recombination but allows expression of added genetic sequences in place of the endogenous gene. This approach allows the test of more subtle mutations than is allowed by a simple knock-out.

[0032] The term "epitope" describes a region on a macromolecule which is recognized by an antibody. Frequently it is in a short region of primary sequence in a protein and it is generally about 5 to 12 amino acids long (the size of the antigen binding site on an antibody). Carbohydrates, nucleic acids and other macromolecules may be antigens and have epitopes.

#### Detailed Description of the Invention

##### Materials and Methods

[0033] Isolation of the FADS1 and FADS2 cDNAs cDNA fragments corresponding to FADS1 and FADS2 were identified by direct cDNA selection. The cDNA selection was performed essentially as described (Rommens et al. 1993) with only minor modifications. Briefly, total RNA was prepared from human retina and from established human retinal pigment epithelium cell line ARPE-19 (Dunn et al. 1996). Prior to the use as templates for cDNA synthesis the isolated RNAs were separated on a 1.2% agarose gel in the presence of 3-(*N*-morpholino)propanesulfonic acid (MOPS) and formaldehyde to check their integrity (Sambrook et al., 1989).

[0034] RNAs were reverse transcribed using the SUPERSCRIPT™ preamplification system for first strand cDNA synthesis (Gibco, BRL) and the RXGT<sub>12</sub> oligonucleotide primer (5'-CGG AAT TCT CGA GAT CTT TTT TTT TT-3'). After poly(A)-tailing with terminal transferase (United States Biochemical, USB), a cDNA pool was generated by RXGT<sub>12</sub>-primed PCR at 94°C for 1 min; 2 cycles of 94°C, 30 sec; 37°C, 1 min, 72°C, 2 min followed by 22 cycles of 94°C, 30 sec; 58°C, 30 sec and 72°C, 2 min. Prior to hybridization the cDNA pools were pre-annealed to C<sub>0</sub>t-1 DNA (Gibco, BRL) enriched with sonicated LINE1 sequences.

[0035] Genomic PAC clones for cDNA selection were derived from 11q12-q13.1, a region known to contain the gene

underlying Best's vitelliform macular dystrophy (Stöhr et al. 1998). The assembly and orientation of the clones have been described previously (Cooper et al. 1997). Inserts from PAC clones dJ465G21 and dJ139E20 (~1 µg) were isolated by NotI digestion, purified using QIAEXII agarose gel extraction beads (Qiagen) and immobilized on Hybond-N+ membrane filters with an average concentration of 60 ng/mm<sup>2</sup>. The insert filters were subjected to two consecutive rounds of hybridization with a starting mixture of 20 µg of retina and ARPE-19 derived cDNAs. Hybridization time was four days at 58°C in Church hybridization buffer (Church and Gilbert 1984). Filters were washed three times in 2 x SSC/0.1% SDS at room temperature, once each in 0.5 x SSC/0.1% SDS, 0.2 x SSC/0.1% SDS and 0.2 x SSC/0.05% SDS (all at 58°C). A final wash was in 2 x SSC. cDNAs were eluted in distilled H<sub>2</sub>O by incubating for 10 min at 98°C and reamplified by PCR using the RXGT<sub>12</sub> oligonucleotide primer. Four µg of the reamplified cDNAs were used for a second round of hybridization. After two rounds of selection the cDNAs were amplified using the RXGT<sub>12</sub> oligonucleotide primer, digested with EcoRI and cloned into the EcoRI site of pBluescript (Stratagene).

[0036] The selected cDNAs represent segments of the 3'-untranslated region (3'-UTR) of FADS1 (clone IVC4 at FADS1 nucleotide position 3793-4204; clone IVB7 at nucleotide position 3132-3609; done VIIC6 at nucleotide position 2077-2317) (Fig. 3) and of the 3' UTR/coding sequence of FADS2 (done IVB8 at FADS2 nucleotide position 2626-3009; clone TUK8-4B at nucleotide position 753-1508) (Fig. 4).

[0037] Using the selected clone sequences extensive dbEST database searches were conducted and revealed a large number of additional overlapping expressed sequence tags (ESTs). More than 100 ESTs (e.g. zk09h08, EST177650, yb28c03, ym29b05, yx67h05) were assembled to an overlapping EST contig representing FADS1. The assembled EST sequences contain an open reading frame (ORF) of 1410 bp, with a first potential in-frame translation initiation codon, ATG, starting 79 nucleotides downstream the most 5'end of EST clone zk09h08.r1 (GenBank acc. no. AA029030) (Fig. 1a). A consensus polyadenylation signal, AAUAAA, was identified at nucleotide position 4.182. The mature protein predicted from the ORF consists of 444 amino acid residues resulting in a calculated molecular mass of 52.0 kDa (Fig. 6).

[0038] Another 30 overlapping ESTs (e.g. cp2485.seq, HSC2EA121, EST06759, ym42c04, nc08c05) were found facilitating the assembly of the FADS2 cDNA. The assembled EST sequences contain an open reading frame (ORF) of 1352 bp, with a first potential in-frame translation initiation codon, ATG, starting 21 nucleotides downstream the most 5'end of EST done ub64e01.r1 (GenBank acc. no. AI036465) (Fig. 4). Consensus polyadenylation signals were predicted at nucleotide positions 2.996 and 4.056. The mature FADS2 protein predicted from the ORF consists of 444 amino acid residues resulting in a calculated molecular mass of 52.3 kDa (Fig. 7). Amino acid sequence identity between FADS1 and FADS2 is 62%.

#### Isolation of the FADS3 cDNA

[0039] Additional 30 human EST clones were available to assemble a third individual cDNA, termed FADS3 (e.g. zs84e06, zs84e05, nq23f05, ya49a19, zs86h09). The existence of a third member of the FADS family was confirmed by PCR mapping of FADS1-, FADS2- and FADS3-specific 3'-UTR fragments revealing three distinct gene loci within a 1.4 Mb PAC contig in 11q12-q13.1 (Cooper et al., 1997). The assembled EST sequences contain an open reading frame (ORF) of 1468 bp, with a first potential in-frame translation initiation codon, ATG, starting 134 nucleotides downstream the most 5'end of EST clone qa99d06.s1 (GenBank acc. no. AI123992) (Fig. 5). The mature protein predicted from the ORF consists of 445 amino acid residues resulting in a calculated molecular mass of 51.2 kDa (Fig. 8). The 3'-UTR of the FADS3 cDNA is represented by several EST clones (e.g. zs86h09.s1, AA279632). A potential polyadenylation signal, AUUAAA, is present at cDNA nucleotide position 1.757 and may be functional as AAUAAA is the most common natural variant of the consensus polyadenylation signal AAUAAA (Fig. 5) (Sheets et al., 1990).

[0040] Amino acid sequence identities between FADS1 and FADS3 as well as between FADS2 and FADS3 are 52% and 63%, respectively. All EST sequences in the dbEST databases could be aligned to one of the three cDNAs, FADS1, FADS2, and FADS3. This suggests that there are no additional members of the FADS family in the human genome.

#### Northern blot analysis

[0041] Northern blot analysis was performed either with total RNA isolated using the guanidinium thiocyanate method (Chomczynski and Sacchi 1987) or with commercially available multiple tissue Northern (MTN) blots purchased from Clontech Laboratories Inc. (Palo Alto, CA). Each lane of the total RNA blot contained 12 µg of total RNA from lung, cerebellum, uterus, retina, liver, heart, RPE cell line ARPE-19, RPE tissue, lymphocytes and was electrophoretically separated in the presence of formaldehyde. The MTN blots were prepared from poly(A)<sup>+</sup> RNA isolated from human heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. Inserts of clones IVC4, IVB7 (FADS1), IVB8 (FADS2) and of the 362 bp PCR product F3/R (5'-ACAGCTTCCCCAATTCTC-3'/5'-GGCCTCAGCTACGAAGT-GAAG-3') (FADS3) derived from the 3'-UTRs of the respective genes were used for filter hybridization at 65°C in 0.5 mM sodium phosphate buffer, pH 7.2; 7% SDS, 1 mM EDTA at 65°C (Church and Gilbert 1984).

[0042] The three genes are ubiquitously expressed and appear to have similar expression levels in all tissues analyzed. FADS1 revealed a transcript size of 4.0 kb while FADS2 revealed a similar sized transcript of 4.0 kb in addition to a smaller transcript of approximately 3.1 kb. The two FADS2 variants may be due to differential usage of polyadenylation signals (see above). Finally, FADS3 is represented by two transcripts of 1.75 kb and 1.25 kb in size. While the former is in agreement with the usage of the variant polyadenylation signal identified at position 1738 of the cDNA, the small size of the latter transcript can not be explained at present and it does not appear to be due to a differential usage of polyadenylation signals. Possibly, differential splicing and/or exon skipping may be involved in the generation of the variant transcript. However, there is no evidence from cDNA cloning or EST contig assembly to support this possibility.

10 Comparison with other desaturases

[0043] Local sequence alignments of the deduced amino acid sequences of FADS1, FADS2, and FADS3 with known proteins or protein motifs were done using SwissProt (<http://www.ncbi.nlm.nih.gov/cgi-bin/Blast/nph-blast?Jform=0>) and the BLASTP and BEAUTY programs at Baylor College of Medicine (<http://dot.imgen.bcm.edu:9331/seq-search/protein-search.html>). Amino acid sequence alignments were performed using the CLUSTALW multiple alignment program at [http://pbil.ibcp.fr/NPSA/npsa\\_clustalw.html](http://pbil.ibcp.fr/NPSA/npsa_clustalw.html). Phylogenetic tree assembly was done using the TREECON software Version 1.3b available at <http://bioc-www.uia.ac.be/u/yvdp/index.html>.

[0044] Overall amino acid identities to known desaturases were found to be in the range of 22% - 27% (Fig. 1). Phylogenetic tree construction revealed a genetic relationship of FADS1, FADS2, and FADS3 to the Δ5-, Δ6- and Δ8-desaturases with some distance to the Δ9-desaturases (Fig. 2). From these analyses it becomes obvious that sequence identity by itself is not a predictor of a specific desaturase activity. For example, Δ5- and Δ6-desaturases from *C. elegans* demonstrate a higher sequence identity to each other than to the Δ6-desaturases from other species. We therefore conclude that based on simple sequence comparisons it is not feasible to determine the specific functions of FADS1, FADS2, and FADS3. This will be done by transgene expression of the three desaturases combined with gas chromatography-mass spectrometry.

[0045] Hydropathy plots of FADS1, FADS2, and FADS3 indicate two hydrophobic sequences predicted to represent transmembrane-spanning domains similar to other desaturases identified thus far (Fig. 1) (reviewed in Sperling et al. 1995).

30 cDNA amplification of FADS1, FADS2, and FADS3

[0046] The coding sequences of the three genes are amplified in overlapping fragments by performing RT-PCR using oligonucleotide primer pairs derived from the respective cDNA sequences:

35 (1) FADS1 (Fig. 9 and SEQ ID NOS. 7-12)

[0047] Sense primer TU12-R5 (5'-CGCCTGACAGCCCCTGCT-3') at cDNA position 31-48 in combination with anti-sense primer TU12-F10 (5'-CAGGTGGCCAATCACAAAT-3') at cDNA position 671-690 results in a product of 660 bp; sense primer TU12-R7 (5'-CTCAAAGTGGAACCATCTGCTA-3') at cDNA position 645-666 in combination with antisense primer TU12-F9 (5'-GGAAACCCAGTCCATGTTCC-3') at cDNA position 1130-1149 results in a product of 505 bp; sense primer TU12-R6 (5'-CCTGGGCCTTTCTTCATAGT-3') at cDNA position 1035-1055 in combination with antisense primer TU12-F5 (5'-CTCAAGCTCCCCTCTGCCT-3') at cDNA position 1465-1483 results in a product of 449 bp.

45 (2) FADS2 (Fig. 9 and SEQ ID NOS. 13-18)

[0048] Sense primer TU13-R4 (5'-TCAGAACGCATAACCTGCGC-3') at cDNA position 98-116 in combination with antisense primer TU13-F7 (5'-CCAGTTCACCAATCAGCAGG-3') at cDNA position 284-303 results in a product of 206 bp; sense primer TU13-R3 (5'-CCCTGCTGATTGGTGAAC-3') at cDNA position 282-301 in combination with anti-sense primer TU13-F4 (5'-TGTAGGGCAGGTATTCAGC-3') at cDNA position 779-798 results in a product of 517 bp; sense primer TU13-R2 (5'-AGCCCACATCGAGTACGGCAA-3') at cDNA position 754-772 in combination with antisense primer TU13-F1 (5'-CCTCAGAACAAAAGCCCAC-3') at cDNA position 1416-1435 results in a product of 682 bp.

(3) FADS3 (Fig. 9 and SEQ ID NOS. 19-22)

[0049] Sense primer TU19-R2 (5'-TCTTGCTCGGACCTCGGC-3') at LLCDL3 cDNA position 81-98 in combination with antisense primer TU19-F2 (5'-GTGATCCACACGAACCAGTG-3') at cDNA position 1130-1149 position results in a product of 1069 bp; sens primer TU19-R3 (5'-GAAGAACCCAGCCAGGATG-3') at cDNA position 428-446 in com-

EP 1 035 207 A1

bination with antisense primer TU19-F3 (5'-ACAGCTTCCCCAATTCTC-3') at cDNA position 1709-1728 results in a product of 1301 bp.

Short description of Figures

5

[0050]

Fig. 1 Comparison of putative amino acid sequences from FADS1, FADS2, FADS3, *Borago officinalis*, *Helianthus annuus* and human cytochrome b5. Arrowheads indicate eight invariant amino acid residues typical for the cytochrome b5 domain. Two potential transmembrane domains are boxed. Three histidine motifs HX<sub>2(3)</sub>[XH]H that are conserved within the desaturase family are hatched.

Fig. 2 Phylogenetic tree of fatty acid desaturases.

Fig. 3 (SEQ ID NO. 1) shows the nucleotide sequence of the FADS1 cDNA

Fig. 4 (SEQ ID NO. 2) shows the nucleotide sequence of the FADS2 cDNA

Fig. 5 (SEQ ID NO. 3) shows the nucleotide sequence of the FADS3 cDNA

Fig. 6 (SEQ ID NO. 4) shows the putative amino acid sequence of the predicted FADS1 protein

Fig. 7 (SEQ ID NO. 5) shows the putative amino acid sequence of the predicted FADS2 protein

Fig. 8 (SEQ ID NO. 6) shows the putative amino acid sequence of the predicted FADS3 protein

Fig. 9 (SEQ ID NOS. 7-22) shows the oligonucleotide PCR primers utilized to amplify the FADS1, FADS2, FADS3 cDNA, respectively.

References

25 [0051] Ariza-Ariza R, Mestanza-Peralta M, Cardiel MH. Omega-3 fatty acids in rheumatoid arthritis: an overview. Semin Arthritis Rheum 27: 366-370 (1998)

30 Calder PC. Immunoregulatory and anti-inflammatory effects of n-3 polyunsaturated fatty acids. Braz J Med Biol Res 31: 467-490 (1998)

Chi Y, Gupta RK. Alterations in membrane fatty acid unsaturation and chain length in hypertension as observed by <sup>1</sup>H NMR spectroscopy. Am J Hypertens 11: 340-348 (1998)

35 Chomczynski, P. and Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162: 156-159 (1987)

Church, G.M., and W. Gilbert. Genomic sequencing. Proc Natl Acad Sci USA 81:1991-1995 (1984)

Cook HW. Fatty acid desaturation and chain elongation in eucaryotes. In: Biochemistry of lipids, lipoproteins and membranes, Vance DE & Vance JE (eds), Elsevier Amsterdam London, New York, pp.141-169 (1991)

40 Cooper P, Nowak NJ, Higgins MJ, Simpson SA, Stöhr H, Marquardt A, Weber BHF, Gerhard DS, deJong P, Shows TB. A sequence ready high resolution physical map of the Best's macular dystrophy gene region in 11q12-q13. Genomics 41: 185-192 (1997)

45 Delton-Vandenbrouke I, Grammas P, Anderson RE. Polyunsaturated fatty acid metabolism in retinal and cerebral microvascular endothelial cells. J Lipid Res 38:147-159(1997)

Dunn KC, Aotaki-Keen AE, Putkey FR, Hjelmeland LM. ARPE-19, a human retinal pigment epithelial cell line with differentiated properties. Exp Eye Res 62: 155-169 (1996)

50 Fan YY, Chapkin RS. Importance of dietary gamma-linolenic acid in human health and nutrition. J Nutr 128: 1411-1414 (1998)

Grattan C, Burton JL, Manku M, Stewart C, Horrobin DF. Essential-fatty-acid metabolites in plasma phospholipids in patients with ichthyosis vulgaris, acne vulgaris and psoriasis. Clin Exp Dermatol 15:174-176 (1990)

Grimble RF, Tappia PS. Modulation of pro-inflammatory cytokine biology by unsaturated fatty acids. Z Ernährungsswiss 37 (Suppl 1): 57-65 (1998)

- 5 Guiard B, Lederer F. The "cytochrome b5 fold": structure of a novel protein superfamily. *J Mol Biol* 135: 639-50 (1979)
- 10 Hodge L, Salome CM, Hughes JM, Liu-Brennan D, Rimmer J, Allman M, Pang D, Armour C, Woolcock AJ. Effect of dietary intake of omega-3 and omega-6 fatty acids on severity of asthma in children. *Eur Respir J* 11: 361-536 (1998)
- 15 Horrobin DF. Abnormal membrane concentrations of 20 and 22-carbon essential fatty acids: a common link between risk factors and coronary and peripheral vascular disease? *Prostaglandins Leukot Essent Fatty Acids* 53: 385-396 (1995)
- 20 Horrobin DF. Essential fatty acids in clinical dermatology. *J Am Acad Dermatol* 20: 1045-1053 (1989)
- 25 Horrobin DF. Fatty acid metabolism in health and disease: the role of Δ6-desaturase. *Am J Clin Nutr* 57(5 Suppl): 732S-737S (1993)
- 30 James MJ, Cleland LG. Dietary n-3 fatty acids and therapy for rheumatoid arthritis. *Semin Arthritis Rheum* 27: 85-97 (1997)
- 35 Jiang WG, Bryce RP, Horrobin DF, Mansel RE. gamma-Linolenic acid blocks cell cycle progression by regulating phosphorylation of p27kip1 and p57kip2 and their interactions with other cycle regulators in cancer cells. *Int J Oncol* 13: 611-617 (1998b)
- 40 Jiang WG, Hiscox S, Bryce RP, Horrobin DF, Mansel RE. The effects of n-6 polyunsaturated fatty acids on the expression of nm-23 in human cancer cells. *Br J Cancer* 77: 731-738 (1998a)
- 45 Jiang WG, Hiscox S, Hallett MB, Horrobin DF, Mansel RE, Puntis MC. Regulation of the expression of E-cadherin on human cancer cells by gamma-linolenic acid (GLA). *Cancer Res* 55: 5043-5048 (1995)
- 50 Jiang WG, Hiscox S, Horrobin DF, Bryce RP, Mansel RE. Gamma linolenic acid regulates expression of maspin and the motility of cancer cells. *Biochem Biophys Res Commun* 237: 639-644 (1997b)
- 55 Jiang WG, Singhrao SK, Hiscox S, Hallett MB, Bryce RP, Horrobin DF, Puntis MC, Mansel RE. Regulation of desmosomal cell adhesion in human tumor cells by polyunsaturated fatty acids. *Clin Exp Metastasis* 15: 593-602 (1997a)
- 60 Leichsenring M, Kochsieck U, Paul K. (n-6)-Fatty acids in plasma lipids of children with atopic bronchial asthma. *Pediatr Allergy Immunol* 6: 209-212 (1995)
- 65 Moore SA, Yoder E, Murphy S, Dutton GR, Spector AA. Astrocytes, not neurons, produce docosahexaenoic acid (22:6 omega-3) and arachidonic acid (20:4 omega-6). *J Neurochem* 56: 518-524 (1991)
- 70 Mori Y, Murakawa Y, Katoh S, Hata S, Yokoyama J, Tajima N, Ikeda Y, Nobukata H, Ishikawa T, Shibutani Y. Influence of highly purified eicosapentaenoic acid ethyl ester on insulin resistance in the Otsuka Long-Evans Tokushima Fatty rat, a model of spontaneous non-insulin-dependent diabetes mellitus. *Metabolism* 46: 1458-1464 (1997)
- 75 Reddy VV, Kupfer D, Caspi E. Mechanism of C-5 double bond introduction in the biosynthesis of cholesterol by rat liver microsomes. *J Biol Chem* 252: 2797-2801 (1977)
- 80 Rommens JM, Lin B, Hutchinson GB, Andrew SE, Goldberg YP, Glaves ML, Graham R, Lai V, McArthur J, Nasir J, Theilmann J, McDonald H, Kalchman M, Clarke LA, Schappert K, Hayden MR. A transcription map of the region containing the Huntington disease gene. *Hum Mol Genet* 2: 901-907 (1993)
- 85 Russo C, Olivieri O, Girelli D, Guarini P, Pasqualini R, Azzini M, Corrocher R. Increased membrane ratios of metabolite to precursor fatty acid in essential hypertension. *Hypertension* 29: 1058-1063 (1997)
- 90 Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor

Laboratory Press, USA (1989)

Shanklin J, Whittle E, Fox BG. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-CoA desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. Biochem 33: 12787-12794 (1994)

Shanklin J, Whittle EJ, Fox BG. Membrane bound desaturases and hydroxylases: Structure function studies, in *Plant lipid metabolism*, Kader JC, Mazliak P (eds), Kluwer Academic Publishers, Netherlands, pp.18-20 (1995).

Sheets MD, Ogg SC, Wickens MP. Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. Nucl Acid Res 18: 5799-5805 (1990)

Singer SJ, Nicolson GL. The fluid mosaic model of the structure of cell membranes. Science 175: 720-731 (1972)

Sperling P, Schmidt H, Heinz E. A cytochrome b5-containing fusion protein similar to plant acyl lipid desaturases. Eur J Biochem 232: 798-805 (1995)

Stöhr H, Marquardt A, Rivera A, Cooper PR, Nowak NJ, Shows TB, Gerhard DS, Weber BHF. A gene map of the Best's vitelliform macular dystrophy region in chromosome 11q12-q13.1. Genome Res 8: 48-56 (1998)

Strittmatter P, Spatz L, Corcoran D, Rogers MJ, Setlow B, Redline R. Purification and properties of rat liver microsomal stearyl coenzyme A desaturase. Proc Natl Acad Sci USA 71: 4565-4569 (1974)

Stubbs CD, Smith AD. The modification of mammalian membrane polyunsaturated fatty acid composition in relation to membrane fluidity and function. Biochim Biophys Acta 779: 89-137 (1984)

Villani F, Comazzi R, De Maria P, Galimberti M. Effect of dietary supplementation with polyunsaturated fatty acids on bronchial hyperreactivity in subjects with seasonal asthma. Respiration 65: 265-269 (1998)

Wang N, Anderson RE. Synthesis of docosahexaenoic acid by retina and retinal pigment epithelium. Biochemistry 32: 13703-13709 (1993)

35

40

45

50

55

Annex to the application documents - subsequently filed sequences listing

[0052]

5

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 10 (i) APPLICANT:  
 (A) NAME: MultiGen Biotech GmbH  
 (B) STREET: Am Hubland  
 (C) CITY: Wuerzburg  
 (D) STATE: -  
 (E) COUNTRY: Germany  
 15 (F) POSTAL CODE (ZIP): 97074  
 (G) TELEPHONE: 0931-7058-4340  
 (H) TELEFAX: 0931-7058-4355  
 (I) TELEX: -
- (ii) TITLE OF INVENTION: cDNA molecules of the members of a gene family  
 20 encoding human fatty acid desaturases and their use in  
 diagnosis and therapy
- (iii) NUMBER OF SEQUENCES: 22
- (iv) COMPUTER READABLE FORM:  
 (A) MEDIUM TYPE: Floppy disk  
 25 (B) COMPUTER: IBM PC compatible  
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

- 30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 444 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: protein
- (ix) FEATURE:  
 (A) NAME/KEY: Protein  
 (B) LOCATION:1..444

## (2) INFORMATION FOR SEQ ID NO: 1:

- 40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4204 base pairs  
 (B) TYPE: nucleic acid  
 45 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:  
 (A) NAME/KEY: exon  
 (B) LOCATION:1..4204

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

	CACTCCTGGA	GCCCCGGAC	CCCGAGCACG	CGCCTGACAG	CCCCTGCTGG	CCCCGGCGCG	60
5	GGCGTCGCCA	GGCCAGCTAT	GGCCCCCGAC	CCGGTGGCCG	CCGAGACCGC		
	GGCTCAGGGA	120					
	CCTACCCCGC	GCTACTTCAC	CTGGGACGAG	GTGGCCCAGC	GCTCAGGGTG		
	CGAGGAGCGG	180					
10	TGGCTAGTGA	TCGACCCTAA	GGTGTACAAC	ATCAGCGAGT	TCACCCGCCG		
	GCATCCAGGG	240					
	GGCTCCCGGG	TCATCAGCCA	CTACGCCGGG	CAGGATGCCA	CGGATCCCTT		
15	TGTGGCCTTC	300					
	CACATCAACA	AGGGCCTTGT	GAAGAACTAT	ATGAACTCTC	TCCTGATTGG		
	AGAACTGTCT	360					
	CCAGAGCAGC	CCAGCTTGA	GCCCACCAAG	AATAAAGAGC	TGACAGATGA		
20	GTTCCGGGAG	420					
	CTGCGGGCCA	CAGTGGAGCG	GATGGGGCTC	ATGAAGGCCA	ACCATGTCTT		
	CTTCCTGCTG	480					
	TACCTGCTGC	ACATCTTGCT	GCTGGATGGT	GCAGCCTGGC	TCACCCTTTG		
25	GGTCTTTGGG	540					
	ACGTCCCTTT	TGCCCTTCT	CCTCTGTGCG	GTGCTGCTCA	GTGCAGTTCA		
	GGCCCAGGCT	600					
	GGCTGGCTGC	AGCATGACTT	TGGGCACCTG	TCGGTCTTCA	GCACCTCAAA		
30	GTGGAACCAT	660					
	CTGCTACATC	ATTTTGTGAT	TGGCCACCTG	AAGGGGGCCC	CCGCCAGTTG		
	GTGGAACCAC	720					
	ATGCACTTCC	AGCACCATGC	CAAGCCCAAC	TGCTTCCGCA	AAGACCCAGA		
35	CATCAACATG	780					
	CATCCCTTCT	TCTTTCGCTT	GGGAAGATC	CTCTCTGTGG	AGCTTGGGAA		
	ACAGAAGAAA	840					
	AAATATATGC	CGTACAACCA	CCAGCACAAA	TACTTCTTCC	TAATTGGGCC		
40	CCCAGCCTTG	900					
	CTGCCTCTCT	ACTTCAGTG	GTATATTTTC	TATTTTGTGA	TCCAGCGAAA		
	GAAGTGGGTG	960					
	GACTTGGCCT	GGATGATTAC	CTTCTACGTC	CGCTTCTTCC	TCACCTATGT		
45	GCCACTATTG	1020					
	GGGCTGAAAG	CCTTCCCTGGG	CCTTTCTTC	ATAGTCAGGT	TCCTGGAAAG		
	CAACTGGTTT	1080					
	GTGTGGGTGA	CACAGATGAA	CCATATTCCC	ATGCACATTG	ATCATGACCG		
50	GAACATGGAC	1140					
	TGGGTTTCCA	CCCAGCTCCA	GGCCACATGC	AATGTCCACA	AGTCTGCCTT		
	CAATGACTGG	1200					
	TTCAGTGGAC	ACCTCAACTT	CCAGATTGAG	CACCATCTTT	TTCCCACGAT		
55	GCCTCGACAC	1260					

5           AATTACCA AAGTGGCTCC CCTGGTGCAG TCCTTGTGTG CCAAGCATGG  
 CATAGAGTAC   1320  
 CAGTCCAAGC CCCTGCTGTC AGCCTCGCC GACATCATCC ACTCACTAAA  
 GGAGTCAGGG  1380  
 CAGCTCTGGC TAGATGCCCTA TCTTCACCAA TAACAACAGC CACCCCTGCC  
 AGTCTGGAAG  1440  
 10          AAGAGGGAGGA AGACTCTGGA GCCAAGGCAG AGGGGAGCTT GAGGGACAAT  
 GCCACTATAG  1500  
 TTTAATACTC AGAGGGGGTT GGGTTTGGGG ACATAAAGCC TCTGACTCAA  
 ACTCCTCCCT  1560  
 15          TTTATCTTCT AGCCACAGTT CTAAGACCCA AAGTGGGGGG TGGACACAGA  
 AGTCCCTAGG  1620  
 AGGGAAGGAG CTGTTGGGGC AGGGGTGTAA ATTATTTCTT TTTTCTAGTT  
 TGGCACATGC  1680  
 20          AGGTAGTTGG TGAACAGAGA GAACCAGGAG GTAAACAGAA GAGGAGGGAC  
 CTACTGAACC  1740  
 CAGAGTCAGG AAGAGATTAA ACACTAAAAT TCCACTCATG CGGGCGTGG  
 TGCAACGCC  1800  
 25          TGTAATCCCCA GCTACCCAGG AGGCTGAGGC AGGAGAACCG CTTGAACCGG  
 GGAGGTGGAG  1860  
 GTTGCAGTGA GCTGAGATCA CGCCATTGTA CTCCAGCCTG GGCACAGAG  
 CAAGACTCCA  1920  
 30          TTTCAAAAAAA AAAAAAAAAAA AAAAATCC ACTCATATAA AAGGTGAGCT  
 CAGCTCACTG  1980  
 GTCATTTCT CAGTGGCTTC TCCATCCTCA TTTGCAAACC TCAGAGGGAT  
 AAGGCAGTTG  2040  
 AACCTGATGA GCAAGAATTAA TAACAGCAAG GAAACATTAA TGCTTAGAAT  
 TCTGAGATCC  2100  
 35          AGCACAACTC AGTCTGTGG AGCTCAGCTC GCTGCCAGG GATAGGTATG  
 ACCTATGTCT  2160  
 GCCTTAGGCT GCTGGGAGAT GCCATTCTCC AGTTTCAGAA GCAGGCAGGG  
 CAAAGGTCAA  2220  
 40          GACTGTGGTA TTGGGGCTTT TTGGCTCTGA AGGATCCTGG AACCACTGAT  
 TTTGGTTTAT  2280  
 TCCCTCCAGG GTCTAAAGAG ACAAGAGGT GCTAGCTCTT ACCAAAACAG  
 ATGGTAGAGA  2340  
 45          GAGTTGCTGG CTATTTAAAA AGCTCTTCA TCTTTAATT CACCTCTCT  
 TTTCACCTCT  2400  
 TIAACCACTC CTCAGGAACA GAACACTTCT AGGACTGGGG GTCTTTAGC  
 TCCATAAGCA  2460  
 50          AGTGAGCAGA TGGGACAAGT TAGTCTTTC TCCCTAGAAA CAAAGGGAT  
 GCCCAGTGGT  2520

TTCCCTTTGC TTCCCAACCT AAAATTCAA GTTTAATAAA ATAGCAATTA  
 GCAGAAAGTGA 2580

5 CCAAATTGGG AGATAATTAT CAGTCATGAG GAAAGACACA GATTCGGTC  
 ATAAAGAATG 2640

TAAGGGCTAT AAGTAGAAC TTTCTATAAC CTAAATGATG TTATAGAATT  
 ATTTTGAGC 2700

10 AGGAGCAGAA AGATTAATAA TGATCACTTC ATACTTCTAA ATCAGAAATA  
 GGAAGATTAA 2760

AACCACAGAA CAGTTTGTA TTTCTATTGC TGGTAGCTAG GTATCTACT  
 CTGTCCACTC 2820

15 TTGTTCAAGT ATCTAACTCT TCTGGAAACC AAATAGGCTT TAGAAGAGAT  
 TATCCTATAT 2880

TCCTATCAGT ATAATACTAA AATGTAACCT TTTAATCATC TGGTTTTAA  
 AAGATAAACAA 2940

20 GTTTAGCCC TCTCTCCAGA GAGCAAACAT AGGAATATGA CTCAGGAGCC  
 TCCTAGGGCT 3000

TATCATCAGC CCTCACACCC GCTTCCCCCT CCAACCCACA GCCTTGCTT  
 CCAGGTGGCA 3060

25 GGATTACTAC TTTGCCTCTT CAGCAGCATE TACTCTAGGC ATATTGATCA  
 TTTTAGACAC 3120

TGGGAGAAGA GAACCTCAAA CTAGGAGGAA AAGACAGAGC CTCCACTTAG  
 TTTTGGGAGG 3180

30 GGATGGCAGA CAGTCAAGGA GATGAGCGTC CTAAGGCATG TTGGGATAGG  
 GTCAGATGCA 3240

CCACCCATGG AGAGGTTTGT CAACACAAAG ACATGGAAGG TTAGAGGTTT  
 GTCAACAAAA 3300

35 AGACATGGAA GGTTAGGTTT GTCAACACAA AGACATGGAA GATTAGAGGT  
 TTGTCAACAC 3360

AAAGATACAG GAAGAATGGG CTGCAGAAGA TTTAGATGTT TTCCATTG  
 GCACATTTA 3420

40 CTTAGCTGGA GAACTAGGTT TAAAACAGCC TGGGTAGGAA AATTAGAAGC  
 AAGCTGGATG 3480

CAGTGGCTCA TGCCTGTAAT CCCAACACTT TTGGGAGGTC CAGGCAGGAG  
 GATCACTTGG 3540

45 GCCCAGGGAGG TCAAGCCTGC AGCGAGCTGA GATCACACCA CTGCACTCCA  
 GCCTGGGTG 3600

ATAGAACAAAG ACCCTGTCTC AAAAAAAAAA AAAAACAAACA AAAACTTAA  
 ATTGAGGAGT 3660

50 TGTACCTCCA TTGGCTTCCT CACTCCAAAA TAGGTGCTGA TCCTTCCTAT  
 TCCTATTCTT 3720

TGCCACCTTT TGGGTGTGGT GTCACCAGCC TGTTTAGCCA AGTAGCTTG  
 GGCATAGGCT 3780

5 GCCCAATCTG AGCAAACACC AGTGAGGCTC TATTGAGCAA GACCAAGTCC  
TCAAAGCACC 3840

10 TGAACCACTG TGGCCTTCTC AGCCTACAGC AGTGTGGTCT CTTACATGGC  
CACAAAGGGA 3900

CACACAGTGA CAAAAGGCTC GGAATGTTAC AATGGTAAAA TGAGTGATCT  
CAAATCCACT 3960

15 GACAGATATA AAATAGGCTT AGAGAGGAAA AGCTGCCTCT GGTCAAGTAG  
ATCATGGCAG 4020

CATGAATTCC AACTCACTTT TTTACGAACT CCAACTTCTA TGTTTATCTT  
TGTTACTTTC 4080

20 ACTTTTTTAC AACCTGCAGA GGCATTTTT AAATCAGGCC CAATATCAGT  
ATTCTTTTG 4140

TGTGTGCCAA TTTGTATC ACATCCCTAT GAAGTTGAAA AATAAAGTTA  
ATTTGACCA 4200

25 AAAG 4204

30

35

40

45

50

55

EP 1 035 207 A1

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:

  - (A) LENGTH: 4089 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..4089

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CGTCACAGTC GGCAGGCAGC ATGGGGAAAGG GAGGGAACCA GGGCGAGGGG  
GCCGCCGAGC 60

GCGAGGTGTC GGTGCCACC TTCAGCTGGG AGGAGATTCA GAAGCATAAC  
CTGCGCACCG 120

ACAGGTGGCT GGTCAATTGAC CGCAAGGTTT ACAACATCAC CAAATGGTCC  
ATCCAGCACC 180

CGGGGGGCCA GCGGGTCATC GGGCACTACG CTGGAGAAGA TGCAACGGAT  
GCCTTCCGCG 240

CCTTCCACCC TGACCTGGAA TTCGTGGGCA AGTTCTTGAA ACCCCTGCTG  
ATTGGTGAAAC 300

TGGCCCCGGA GGAGCCCAGC CAGGACCACG GCAAGAACTC AAAGATCACT  
GAGGACTTCC 360

GGGCCCTGAG GAAGACGGCT GAGGACATGA ACCTGTTCAA GACCAACCAC  
GTGTTCTTCC 420

TCCTCTCTCTTGGGCCCCACATC ATCGCCCTGG AGAGGATTTGC ATGGTTCACT  
GTCTTTACT 480

TGGGATAGG CTGGATCCCT ACCTCTATCA CGGCCCTTGT CCTTGCTACC  
TCTCAGGCC 540

AACCTGGATG CCCTCCACAT GATTAGGGCC ACCCTGCTGT CTACAGAAAA  
CCCCAAGTGGAA 600

AACTGGTGGGA 660

CCCGATGTGA 720

AAGAAGAAGC 780

CCGCCGCTGC 840

AAGAACTGGG 900

TGGACCTGGC CTGGGGCCGTC AGCTACTACA TCCGGTTCTT CATCACCTAC  
 ATCCCTTTCT 960

5 ACGGCATCCT GGGAGCCCTC CTTTCCTCA ACTTCATCAG GTTCCTGGAG  
 AGCCACTGGT 1020

TTGTGTGGGT CACACAGATG AATCACATCG TCATGGAGAT TGACCAGGAG  
 GCCTACCGTG 1080

10 ACTGGTTCAAG TAGGCCAGCTG ACAGCCACCT GCAACGTGGA GCAGTCCTTC  
 TTCAACGACT 1140

GGTTCACTGG ACACCTTAAC TTCCAGATTG AGCACCACCT CTTCCCCACC  
 ATGCCCGGC 1200

15 ACAACTTACA CAAGATCGCC CCGCTGGTGA AGTCTCTATG TGCCAAGCAT  
 GGCATTGAAT 1260

ACCAGGGAGAA GCCGCTACTG AGGGCCCTGC TGGACATCAT CAGGTCCCTG  
 AAGAAGTCTG 1320

20 GGAAGCTGTG GCTGGACGCC TACCTTCACA AATGAAGCCA CAGCCCCCGG  
 GACACCGTGG 1380

GGAAGGGGTG CAGGTGGGGT GATGCCAGA GGAATGATGG GCTTTGTTC  
 TGAGGGGTGT 1440

25 CCGAGAGGCT GGTGTATGCA CTGCTCACGG ACCCCATGTT GGATCTTTCT  
 CCCTTTCTCC 1500

TCTCCTTTCT CTCCTCACAT CTCCCCATA GCACCCCTGCC CTCATGGAC  
 CTGCCCTCCC 1560

30 TCAGCCGTCA GCCATCAGCC ATGGCCCTCC CAGTGCCTCC TAGCCCCCTTC  
 TTCCAAGGAG 1620

CAGAGAGGTG GCCACCGGGG GTGGCTCTGT CCTACCTCCA CTCTCTGCC  
 CTAAAGATGG 1680

35 GAGGAGACCA GCGGTCCATG GGTCTGGCCT GTGAGTCTCC CCTTGCAGCC  
 TGGTCACTAG 1740

GCATCACCCCC CGCTTGGTT CTTCAGATGC TCTTGGGGTT CATAGGGGCA  
 GGTCTAGTC 1800

40 GGGCAGGGCC CCTGACCCCTC CCGGCCTGGC TTCACCTCTCC CTGACGGCTG  
 CCATTGGTCC 1860

ACCCTTCAT AGAGAGGCCT GCTTTGTTAC AAAGCTCGGG TCTCCCTCCT  
 GCAGCTCGGT 1920

45 TAAGTACCCG AGGCCTCTCT TAAGATGTCC AGGGCCCCAG GCCCGCGGGC  
 ACAGCCAGCC 1980

CAAACCTTGG GCCCTGGAAG AGTCCTCCAC CCCATCACTA GAGTGCTCTG  
 ACCCTGGGCT 2040

50 TTCACGGGCC CCATTCCACC GCCTCCCCAA CTTGAGCCTG TGACCTTGGG  
 ACCAAAGGGG 2100

GAGTCCCTCG TCTCTTGTGA CTCAGCAGAG GCAGTGGCCA CGTCAGGGGA  
 GGGGCCGGCT 2160

55

## EP 1 035 207 A1

5 GGCCTGGAGG CTCAGCCCCAC CCTCCAGCTT TTCCCTCAGGG TGTCCTGAGG  
 TCCAAGATTC 2220  
 TGGAGCAATC TGACCCCTTCT CCAAAGGCTC TGTTATCAGC TGGGCAGTGC  
 CAGCCAATCC 2280  
 CTGGCCATTG GGCCCCAGGG GACGTGGGCC CTGCAGGCTG CAGGAGGGCA  
 CTGGAGCTGG 2340  
 10 GAGGTCTCGT CCCAGCCCTC CCCATCTCGG GGCTGCTGTG TGGACGGGCC  
 TGCCTCAGGC 2400  
 ACTCTCTGT CTGAACCTGC CCTTAACCTGTG TTTAACCTGT TGCTCCAGGA  
 TGCATTCTGA 2460  
 15 TAGGAGGGGG CGGCAGGGCT GGGCCTTGTG ACAATCTGCC TTTCACCAACA  
 TGGCCTTGCC 2520  
 TCGGTGGCCC TGACTGTCAAG GGAGGGCCAG GGAGGCAGAG CGGGAGGGAG  
 TCTCAGGAGG 2580  
 20 AGGCTGCCCT GAGGGGCTGG GGAGGGGTA CCTCATGAGG ACCAGGGTGG  
 AGCTGAGAAG 2640  
 AGGAGGGAGGT GGGGGCTGGA GGTGCTGGTA GCTGAGGGGA CGGGCAAGTG  
 AGAGGGGAGG 2700  
 25 GAGGGAAGTC CTGGGAGGAT CCTGAGCTGC TGTTGCAGTC TAACCCACTA  
 ATCAGTTCTT 2760  
 AGATTCAGGG GAAGGGCAGG CACCAACAAC TCAGAATGGG GGCTTTCGGG  
 GAGGGCCTT 2820  
 30 AGTCCCCCCTA GCTCTAAGCA GCCAGGAGGG ACCTGCATCT AAGCATCTGG  
 GTTGCCATGG 2880  
 CAATGGCATG CCCCCCAGCT ACTGTATGCC CCCGACCCCC GCAGAGGCAG  
 AATGAACCCA 2940  
 35 TAGGGAGCTG ATCGTAATGT TTATCATGTT ACTTCCCCAC CCCTACATTT  
 TTTGAAATAA 3000  
 AATAAGGAAT TTTATTCTCA CTTCCTGTGT TTCCCTGCACG CCAATGCCAG  
 GCCATGGTAT 3060  
 40 TGGGTGATAG ATGAGGCCCT TCTAGCTGGG CCTGGGCACC AGGAGGGGTC  
 CCCATGCTTG 3120  
 CATCTCTCTG TATCCCCCTCC CTCCCTGTG GCCATCCCAC CGGCCTCTCC  
 CTGCTGCCTC 3180  
 45 TGAAATTCTAT TCTGGGGCCC GGAACTTGGT GGAAATGACC CAAAAACATT  
 GGCCCCATCTT 3240  
 CCTCCTCTCA GCAGCCGACC CCAGCCCAAT TCTAAACAG GGCTGAGAGC  
 CACCTCTCAG 3300  
 50 CAGCTGACCC CTACCCAAGG AGGGTGGCAT GGAGGGGCTT GCAGAGACTC  
 TTCCTAACAT 3360  
 CCTCCCCCCC CAGCTGTCTC CCCAAGTGCA ATCTGCCCTC CCATCCCTGG  
 GCCAGCCAGC 3420

55

TTCCACAGAG CGCCAGGCCA AACAGAATTC CTGGCCTCCT TGGAAGGGC  
 TGGAGAAGGC 3480

5 CGGGAGCAGT GGCTCACGCC TGTAATCCA GCACTTTGGG AGGCTGAGGC  
 GGGCAGATCA 3540

CAAAGTCAAG AGATTGAGAC CATCCTGCC AACATGGTGA AACCCCGTCT  
 CTACTAAAAA 3600

10 TACAAAAAATT AGGCCGGGTG CGGTGGCTCA CGCCTGTAAT CCCAGCACTT  
 TGGGAGGCCG 3660

15 AGGCAGGGCAG ATCACGAGGT CAGGAGATCA AGACCATCCT GGCTAACACG  
 GTGAAACCCC 3720

GTCTCTACTA AAAATACAAA AAATTAGCTG GGCGAGGTGG CGGGTGCCTG  
 TAGTCCCAGC 3780

20 TACGTGGGAG GCTGAGGCAA GAGAATGGCG TGAACCCGG CGGGGCAGAG  
 CCTGCAGAGA 3840

GCTGAGATCA CACCACTGTA CTCCAGCCTG GGCGACAGCG AGACTCCGTC  
 TCAAAAAAAA 3900

25 AAAAAAAAAA AATTAGCTGG GCATGGTGGT GCGTGCCTGC AGTCCCAGCT  
 ACTCAGGAGG 3960

CTGAGACGGG AGAATCGCTT GAACCTGGGA GGCAGAGGTT GCAGTGAGCC  
 AAGATCGCTC 4020

30 ACTCCAGCCT AGCGACAGAG TGAGACTCCA TCTCAAATAA ATAAATAAAT  
 TAATTAATT 4080

AATTAATT 4089

35

40

45

50

55

EP 1 035 207 A1

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:  
5 (A) LENGTH: 1757 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

15 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION: 1..1757

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

20 GGC CGCGGCG GCAGGGCGGG GCCGGAGCAG CGGGCGGCGG CGGAGGC GGCGCGGAGC 60

25 GCTCTTCGCT TCCCTCGGGG TCTTGCTCGG ACCTCGGCCA CCGCCTGGGA TCCCCAGGAC 120

TCGTGCGTGC AGCATGGCG GCGTCGGGGA GCCGGGACCG CGGGAGGGAC CCGCGCAGCC 180

25 GGGGCACCG CTGCCCACCT TCTGCTGGGA GCAGATCCGC GCGCACGACC AGCCCAGGCA 240

CAAGTGGCTG GTCATCGAGC GCCGCGTCTA CGACATCAGC CGCTGGGCAC AGCGGCACCC 300

30 AGGGGGCAGC CGCCTCATCG GCCACCACGG CGCTGAGGAC GCCACGGATG CCTTCCGTGC 360

CTTCCATCAA GATCTCAATT TTGTGCGCAA GTTCCTACAG CCCCTGTTGA TTGGAGAGCT 420

35 GGCTCCGGAA GAACCCAGCC AGGATGGACC CCTGAATGCG CAGCTGGTCG AGGACTTCCG 480

AGCCCTGCAC CAGGCAGCCG AGGACATGAA GCTGTTTGAT GCCAGTCCCA CCTTCTTTCG 540

40 TTTCCTACTG GGCCACATCC TGGCCATGGA GGTGCTGGCC TGGCTCCTTA TCTACCTCCT 600

GGGTCCCTGGC TGGGTGCCCA GTGCCCTGGC CGCCTTCATC CTGGCCATCT CTCAGGCTCA 660

45 GTCCTGGTGT CTGCAGCATG ACCTGGGCCA TGCCTCCATC TTCAAGAAGT CCTGGTGGAA 720

CCACGTGGCC CAGAACGTC TGATGGGCA GCTAAAGGGC TTCTCCGCC ACTGGTGGAA 780

50 CTTCCGCCAC TTCCAGCACC ACGCCAAGCC CAACATCTTC CACAAAGACC CAGACGTGAC 840

GGTGGCGCCC GTCTTCCCTCC TGGGGGAGTC ATCCGTGAG TATGGCAAGA AGAAACGCAAG 900

5

10

15

20

25

30

35

40

45

50

55

ATACCTACCC TACAACCAGC AGCACCTGTA CTTCTTCCTG ATCGGCCGC  
 CGCTGCTCAC 960

5 CCTGGTGAAC TTTGAAGTGG AAAATCTGGC GTACATGCTG GTGTGCATGC  
 AGTGGGCGGA 1020

10 TTTGCTCTGG GCCGCCAGCT TCTATGCCCG CTTCTTCTTA TCCTACCTCC  
 CCTTCTACGG 1080

15 CGTCCCTGGG GTGCTGCTCT TCTTTGTTGC TGTCAGGGTC CTGGAAAGCC  
 ACTGGTTCGT 1140

20 GTGGATCACA CAGATGAACC ACATCCCCAA GGAGATCGGC CACGAGAAGC  
 ACCGGGACTG 1200

25 GGTCAGCTCT CAGCTGGCAG CCACCTGCAA CGTGGAGCCC TCACTTTCA  
 CCAACTGGTT 1260

30 CAGCGGGCAC CTCAACTTCC AGATCGAGCA CCACCTCTTC CCCAGGATGC  
 CGAGACACAA 1320

35 CTACAGCCGG GTGGCCCCGC TGGTCAAGTC GCTGTGTGCC AAGCACGGCC  
 TCAGCTACGA 1380

40 AGTGAAGCCC TTCCTCACCG CGCTGGTGGA CATCGTCAGG TCCCTGAAGA  
 AGTCTGGTGA 1440

45 CATCTGGCTG GACGCCCTACC TCCATCAGTG AAGGCAACAC CCAGGGGGC  
 AGAGAAGGGC 1500

50 TCAGGGCACC AGCAACCAAG CCAGCCCCGG CGGGATCGAT ACCCCCACCC  
 CTCCACTGGC 1560

55 CAGCCTGGGG GTGCCCTGCC TGCCCTCCTG GTACTGTTGT CTTCCCCTCG  
 GCCCCCTCAC 1620

60 ATGTGTATTC AGCAGCCCTA TGGCCTTGGC TCTGGGCCTG ATGGGACAGG  
 GGTAGAGGGA 1680

65 AGGTGAGCAT AGCACATTT CCTAGAGCGA GAATTGGGGG AAAGCTGTTA  
 TTTTTATATT 1740

70 AAAATACATT CAGATGT 1757

45

50

55

(2) INFORMATION FOR SEC ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 444 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:  
(A) NAME/KEY: Protein  
(B) LOCATION: 1 444

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4.

Met Ala Pro Asp Pro Val Ala Ala Glu Thr Ala Ala Gln Gly Pro Thr  
1 5 10 15

Pro Arg Tyr Phe Thr Trp Asp Glu Val Ala Gln Arg Ser Gly Cys Glu  
20 25 30

Glu Arg Trp Leu Val Ile Asp Arg Lys Val Tyr Asn Ile Ser Glu Phe  
 35                  40                  45

Thr Arg Arg His Pro Gly Gly Ser Arg Val Ile Ser His Tyr Ala Gly  
50 55 60

Gln Asp Ala Thr Asp Pro Phe Val Ala Phe His Ile Asn Lys Gly Leu  
65 70 75 80

Val Lys Lys Tyr Met Asn Ser Leu Leu Ile Gly Glu Leu Ser Pro Glu  
85 90 95

Gln Pro Ser Phe Glu Pro Thr Lys Asn Lys Glu Leu Thr Asp Glu Phe  
100 105 110

Arg Glu Leu Arg Ala Thr Val Glu Arg Met Gly Leu Met Lys Ala Asn  
115 120 125

His Val Phe Phe Leu Leu Tyr Leu Leu His Ile Leu Leu Leu Asp Gly  
130 135 140

Ala Ala Trp Leu Thr Leu Trp Val Phe Gly Thr Ser Phe Leu Pro Phe  
145 150 155 160

Leu Leu Cys Ala Val Leu Leu Ser Ala Val Gln Ala Gln Ala Gly Trp  
165 170 175

Leu Gln His Asp Phe Gly His Leu Ser Val Phe Ser Thr Ser Lys Trp  
180 185 190

Asn His Leu Leu His His Phe Val Ile Gly His Leu Lys Gly Ala Pro  
195 200 205

Ala Ser Trp Trp Asn His Met His Phe Gln His His Ala Lys Pro Asn  
210 215 220

Cys Phe Arg Lys Asp Pro Asp Ile Asn Met His Pro Phe Phe Phe Ala  
225 230 235 240

## EP 1035 207 A1

	Leu	Gly	Lys	Ile	Leu	Ser	Val	Glu	Leu	Gly	Lys	Gln	Lys	Lys	Tyr		
															255		
5		245															
	Met	Pro	Tyr	Asn	His	His	Gln	His	Lys	Tyr	Phe	Phe	Leu	Ile	Gly	Pro	Pro
															270		
	260								265								
	Ala	Leu	Leu	Pro	Leu	Tyr	Phe	Gln	Trp	Tyr	Ile	Phe	Tyr	Phe	Val	Ile	
10		275							280							285	
	Gln	Arg	Lys	Lys	Trp	Val	Asp	Leu	Ala	Trp	Met	Ile	Thr	Phe	Tyr	Val	
		290							295				300				
15	Arg	Phe	Phe	Leu	Thr	Tyr	Val	Pro	Leu	Leu	Gly	Leu	Lys	Ala	Phe	Leu	
	305							310			315			320			
	Gly	Leu	Phe	Phe	Ile	Val	Arg	Phe	Leu	Glu	Ser	Asn	Trp	Phe	Val	Trp	
		325						330						335			
20	Val	Thr	Gln	Met	Asn	His	Ile	Pro	Met	His	Ile	Asp	His	Asp	Arg	Asn	
		340							345					350			
	Met	Asp	Trp	Val	Ser	Thr	Gln	Leu	Gln	Ala	Thr	Cys	Asn	Val	His	Lys	
		355						360					365				
25	Ser	Ala	Phe	Asn	Asp	Trp	Phe	Ser	Gly	His	Leu	Asn	Phe	Gln	Ile	Glu	
		370						375					380				
	His	His	Leu	Phe	Pro	Thr	Met	Pro	Arg	His	Asn	Tyr	His	Lys	Val	Ala	
30		385						390			395			400			
	Pro	Leu	Val	Gln	Ser	Leu	Cys	Ala	Lys	His	Gly	Ile	Glu	Tyr	Gln	Ser	
														410	415		
35	Lys	Pro	Leu	Leu	Ser	Ala	Phe	Ala	Asp	Ile	Ile	His	Ser	Leu	Lys	Glu	
									420		425			430			
	Ser	Gly	Gln	Leu	Trp	Leu	Asp	Ala	Tyr	Leu	His	Gln					
		435							440								
40																	

45

50

55

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 444 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

10

## (ix) FEATURE:

- (A) NAME/KEY: Protein  
 (B) LOCATION: 1..444

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met	Gly	Lys	Gly	Gly	Asn	Gln	Gly	Glu	Gly	Ala	Ala	Glu	Arg	Glu	Val
1															15

20

Ser	Val	Pro	Thr	Phe	Ser	Trp	Glu	Glu	Ile	Gln	Lys	His	Asn	Leu	Arg
									20						30

Thr	Asp	Arg	Trp	Leu	Val	Ile	Asp	Arg	Lys	Val	Tyr	Asn	Ile	Thr	Lys
									35						45

25

Trp	Ser	Ile	Gln	His	Pro	Gly	Gly	Gln	Arg	Val	Ile	Gly	His	Tyr	Ala
									50						60

30

Gly	Glu	Asp	Ala	Thr	Asp	Ala	Phe	Arg	Ala	Phe	His	Pro	Asp	Leu	Glu
									65						80

Phe	Val	Gly	Lys	Phe	Leu	Lys	Pro	Leu	Leu	Ile	Gly	Glu	Leu	Ala	Pro
									85						95

35

Glu	Glu	Pro	Ser	Gln	Asp	His	Gly	Lys	Asn	Ser	Lys	Ile	Thr	Glu	Asp
								100							110

Phe	Arg	Ala	Ile	Arg	Lys	Thr	Ala	Glu	Asp	Met	Asn	Ile	Phe	Lys	Thr
								115							125

40

Asn	His	Val	Phe	Phe	Leu	Leu	Leu	Leu	Ala	His	Ile	Ile	Ala	Leu	Glu
									130						140

Ser	Ile	Ala	Trp	Phe	Thr	Val	Phe	Tyr	Phe	Gly	Asn	Gly	Trp	Ile	Pro
									145						160

45

Thr	Leu	Ile	Thr	Ala	Phe	Val	Leu	Ala	Thr	Ser	Gln	Ala	Gln	Ala	Gly
									165						175

Trp	Leu	Gln	His	Asp	Tyr	Gly	His	Leu	Ser	Val	Tyr	Arg	Lys	Pro	Lys
								180							190

50

Trp	Asn	His	Leu	Val	His	Lys	Phe	Val	Ile	Gly	His	Leu	Lys	Gly	Ala
									195						205

Ser	Ala	Asn	Trp	Trp	Asn	His	Arg	His	Phe	Gln	His	His	Ala	Lys	Pro
									210						220

55

Asn	Ile	Phe	His	Lys	Asp	Pro	Asp	Val	Asn	Met	Leu	His	Val	Phe	Val
									225						240

## EP 1035 207 A1

	Leu	Gly	Glu	Trp	Gln	Pro	Ile	Glu	Tyr	Gly	Lys	Lys	Lys	Leu	Lys	Tyr
					245					250					255	
5	Leu	Pro	Tyr	Asn	His	Gln	His	Glu	Tyr	Phe	Phe	Leu	Ile	Gly	Pro	Pro
					260				265					270		
	Leu	Leu	Ile	Pro	Met	Tyr	Phe	Gln	Tyr	Gln	Ile	Ile	Met	Thr	Met	Ile
					275				280				285			
10	Val	His	Lys	Asn	Trp	Val	Asp	Leu	Ala	Trp	Ala	Val	Ser	Tyr	Tyr	Ile
					290			295				300				
15	Arg	Phe	Phe	Ile	Thr	Tyr	Ile	Pro	Phe	Tyr	Gly	Ile	Leu	Gly	Ala	Leu
					305			310			315			320		
	Leu	Phe	Leu	Asn	Phe	Ile	Arg	Phe	Leu	Glu	Ser	His	Trp	Phe	Val	Trp
					325			330			335				335	
20	Val	Thr	Gln	Met	Asn	His	Ile	Val	Met	Glu	Ile	Asp	Gln	Glu	Ala	Tyr
					340			345				350				
	Arg	Asp	Trp	Phe	Ser	Ser	Gln	Leu	Thr	Ala	Thr	Cys	Asn	Val	Glu	Gln
					355			360				365				
25	Ser	Phe	Phe	Asn	Asp	Trp	Phe	Ser	Gly	His	Leu	Asn	Phe	Gln	Ile	Glu
					370			375			380					
	His	His	Leu	Phe	Pro	Thr	Met	Pro	Arg	His	Asn	Leu	His	Lys	Ile	Ala
					385			390			395			400		
30	Pro	Leu	Val	Lys	Ser	Leu	Cys	Ala	Lys	His	Gly	Ile	Glu	Tyr	Gln	Glu
					405			410				415				
	Lys	Pro	Leu	Leu	Arg	Ala	Leu	Leu	Asp	Ile	Ile	Arg	Ser	Leu	Lys	Lys
35					420			425			430					
	Ser	Gly	Lys	Leu	Trp	Leu	Asp	Ala	Tyr	Leu	His	Lys				
					435			440								

40

45

50

55

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:

  - (A) LENGTH: 445 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- (ix) FEATURE:  
(A) NAME/KEY: Protein  
(B) LOCATION: 1..495

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Gly Gly Val Gly Glu Pro Gly Pro Arg Glu Gly Pro Ala Gln Pro  
1 5 10 15

Gly Ala Pro Leu Pro Thr Phe Cys Trp Glu Gln Ile Arg Ala His Asp  
20 25 30

Gln Pro Gly Asp Lys Trp Leu Val Ile Glu Arg Arg Val Tyr Asp Ile  
35 40 45

Ser Arg Trp Ala Gln Arg His Pro Gly Gly Ser Arg Leu Ile Gly His  
50 55 60

His Gly Ala Glu Asp Ala Thr Asp Ala Phe Arg Ala Phe His Gln Asp  
65 70 75 80

Leu Asn Phe Val Arg Lys Phe Leu Gln Pro Leu Leu Ile Gly Glu Leu  
85 90 95

Ala Pro Glu Glu Pro Ser Gln Asp Gly Pro Leu Asn Ala Gln Leu Val  
100 105 110

Glu Asp Phe Arg Ala Leu His Gln Ala Ala Glu Asp Met Lys Leu Phe  
115 120 125

Asp Ala Ser Pro Thr Phe Phe Ala Phe Leu Leu Gly His Ile Leu Ala  
130 135 140

Met Glu Val Leu Ala Trp Leu Leu Ile Tyr Leu Leu Gly Pro Gly Trp  
 145                    150                    155                    160

Val Pro Ser Ala Leu Ala Ala Phe Ile Leu Ala Ile Ser Gln Ala Gln  
                   165                 170                 175

Ser Trp Cys Leu Gln His Asp Leu Gly His Ala Ser Ile Phe Lys Lys  
180 185 190

Ser Trp Trp Asn His Val Ala Gln Lys Phe Val Met Gly Gin Leu Lys  
195 200 205

Gly Phe Ser Ala His Trp Trp Asn Phe Arg His Phe Gln His His Ala  
210 215 220

Lys Pro Asn Ile Phe His Lys Asp Pro Asp Val Thr Val Ala Pro Val  
225 230 235 240

EP 1035207 A1

Phe Leu Leu Gly Glu Ser Ser Val Glu Tyr Gly Lys Lys Lys Arg Arg  
245 250 255

5 Tyr Leu Pro Tyr Asn Gln Gln His Leu Tyr Phe Phe Leu Ile Gly Pro  
260 265 270

Pro Leu Leu Thr Leu Val Asn Phe Glu Val Glu Asn Leu Ala Tyr Met  
10 275 280 285

Leu Val Cys Met Gln Trp Ala Asp Leu Leu Trp Ala Ala Ser Phe Tyr  
290 295 300

15 Ala Arg Phe Phe Leu Ser Tyr Leu Pro Phe Tyr Gly Val Pro Gly Val  
305 310 315 320

Leu Leu Phe Phe Val Ala Val Arg Val Leu Glu Ser His Trp Phe Val  
325 330 335

20 Trp Ile Thr Gln Met Asn His Ile Pro Lys Glu Ile Gly His Glu Lys  
340 345 350

His Arg Asp Trp Val Ser Ser Gln Leu Ala Ala Thr Cys Asn Val Glu  
355 360 365

25 Pro Ser Leu Phe Thr Asn Trp Phe Ser Gly His Leu Asn Phe Gln Ile  
370 375 380

Glu His His Leu Phe Pro Arg Met Pro Arg His Asn Tyr Ser Arg Val  
385 390 395 400

30 Ala Pro Leu Val Lys Ser Leu Cys Ala Lys His Gly Leu Ser Tyr Glu  
405 410 415

Val Lys Pro Phe Leu Thr Ala Leu Val Asp Ile Val Arg Ser Leu Lys  
35 420 425 430

Lys Ser Gly Asp Ile Trp Leu Asp Ala Tyr Leu His Gln  
435 440 445

40

45

50

55

(2) INFORMATION FOR SEQ ID NO: 7:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

- 15 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..18

18 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TCGTCCCCGA CAGTCCGC

18

20 (2) INFORMATION FOR SEQ ID NO: 8:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (genomic)

- 35 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..20

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

20 TAAAACACTA ACCGGTGGAC

45 (2) INFORMATION FOR SEQ ID NO: 9:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: DNA (genomic)

- 55 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

ATCGTCTACC AAGGTGAAAC TC

22

(2) INFORMATION FOR SEQ ID NO: 10:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

- 15 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..20

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CCTTGTACCT GACCCAAAGG

20

25 (2) INFORMATION FOR SEQ ID NO: 11:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (genomic)

- 35 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..21

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TGATACTTCT TTTCCGGGTC C

21

45 (2) INFORMATION FOR SEQ ID NO: 12:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: DNA (genomic)

- 55 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TCCGTCTCCC CTCGAACTC

19

(2) INFORMATION FOR SEQ ID NO: 13:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

- 15 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..19

19 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CGCGTCCAAT ACGAAGACT

20 (2) INFORMATION FOR SEQ ID NO: 14:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (genomic)

- 35 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..20

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

20 GGACGACTAA CCACATTGACC

45 (2) INFORMATION FOR SEQ ID NO: 15:

- 40 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: DNA (genomic)

- 55 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

20 TCAAGTGTT AGTCGTCCCC

(2) INFORMATION FOR SEQ ID NO: 16:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

- 15 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION: 1..20

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CGACTTTATG GACGGGATGT

20

25 (2) INFORMATION FOR SEQ ID NO: 17:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (genomic)

- 35 (ix) FEATURE:

- (A) NAME/KEY: exon  
(B) LOCATION: 1..19

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AACGGCATGA GCTACCCGA

19

45 (2) INFORMATION FOR SEQ ID NO: 18:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: DNA (genomic)

- (ix) FEATURE:

- (A) NAME/KEY: exon  
(B) LOCATION: 1..20

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CTACCCGAAA ACAAGACTCC

20

(2) INFORMATION FOR SEQ ID NO: 19:

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

- 15 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CGGCTCCAGG CTCGTTCT

18

20 (2) INFORMATION FOR SEQ ID NO: 20:

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- 30 (ix) FEATURE:

(A) NAME/KEY: exon  
(B) LOCATION:1..20

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GTGACCAAGC ACACCTAGTG

20

40 (2) INFORMATION FOR SEQ ID NO: 21:

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- 50 (ix) FEATURE:

(A) NAME/KEY: exon  
(B) LOCATION:1..19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

55 GTAGGACCGA CCCAAGAAG

19

(2) INFORMATION FOR SEQ ID NO: 22:

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

15 (ix) FEATURE:  
(A) NAME/KEY: exon  
(B) LOCATION:1..20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

20 CTCTTAACCC CCTTTGACA

20

25

30

35

40

45

50

55

## SEQUENCE LISTING

5 <110> MultiGene Biotech GmbH

10 <120> cDNA molecules of the members of gene family encoding  
human fatty acid desaturases and their use in diagnosis  
and therapy

15 <130> MultiGene Anm. 99/002 EP

20 <140> 99104664.0-2105

<141> 1999-03-09

25 <160> 22

30 <170> PatentIn Ver. 2.1

35 <210> 1

<211> 4205

<212> DNA

<213> Unknown Organism

40 <220>

45 <223> Description of Unknown Organism:

50 unknown

55 <400> 1

cactccttgg a cccgcggac cccgagcacg cgcctgacag cccctgctgg cccggcg 60

40 ggcgtcgcca ggcaggctat ggccccggac ccgggtggccg ccgagaccgc ggctcaggga 120

cctaccccgcc gctacttcac ctgggacgag gtggcccagc gtcagggtg cgaggagcgg 180

tggcttagtga tcgaccgtaa ggtgtacaac atcagcagt tcaccccgcc gcattccagg 240

ggctcccggt tcatcagcca ctacggccgg caggatgcca cggatccctt tgtggcc 300

45 cacatcaaca agggccttgt gaagaagtat atgaactctc tcctgattgg agaactgtct 360

ccagagcagc ccagcttga gcccaccaag aataaaagagc tgacagatga gttccgggag 420

ctgcgggcca cagtggagcg gatggggctc atgaaggcca accatgtctt cttccctgctg 480

50 tacctgctgc acatcttgct gctggatggt gcagcctggc tcaccccttg ggtctttggg 540

acgtcccttt tgcccttcct cctctgtgcg gtgctgctca gtgcagttca gggccaggct 600

ggctggctgc agcatgactt tgggcacctg tcggcttca gcacctcaaa gtggaaaccat 660

ctgctacatc attttgtgat tggccacctg aaggggggccc ccgccagttg gtggaaaccac 720

55 atgcacttcc agcaccatgc caagcccaac tgcttccgca aagaccaggaa catcaacatg 780

catcccttct tctttgcctt gggaaagatc ctctctgtgg agcttggaa acagaagaaa 840

aaatataatgc cgtacaacca ccagcacaaa tacttcttcc taattgggcc cccagccttg 900  
 ctgcctctct acttccagtg gtatattttc tattttgtta tccagcggaa gaagtgggtg 960  
 5 gacttggcct ggatgattac cttctacgtc cgcttcttcc tcacttatgt gccactattg 1020  
 gggctgaaag ctttcctggg cctttcttc atagtcaggt tcctggaaag caactggttt 1080  
 gtgtgggtga cacagatgaa ccataatccc atgcacattg atcatgaccg gaacatggac 1140  
 tgggtttcca cccagctcca ggccacatgc aatgtccaca agtctgcctt caatgactgg 1200  
 10 ttcaagtggac acctcaactt ccagattgag caccatctt tccccacgat gcctcgacac 1260  
 aattaccaca aagtggctcc cctggtgcag tccttgcgtg ccaagcatgg catagagtac 1320  
 cagtc当地 cccctgctgtagatgccta tcttccaccaa taacaacagg caccctgccc agtctggaaag 1440  
 15 aagaggagga agactctggaa gccaaggcag aggggagctt gagggacaat gccactata 1500  
 tttaataactc agaggggggtt ggggttgggg acataaaagcc tctgactcaa actcctccct 1560  
 tttatctctc agccacagtt ctaagaccca aagtgggggg tggacacaga agtccctagg 1620  
 20 agggaaaggag ctgttggggc aggggtgtaa attatttcct ttttctagtt tggcacatgc 1680  
 aggttagttgg tgaacagaga gaaccaggag gtaacacagaa gaggaggac ctactgaacc 1740  
 cagagtcaagg aagagatttta acactaaaat tccactcatg ccggggctgg tgacacgccc 1800  
 tgtaatccca gctacccagg aggctgaggc aggagaatcg cttgaaccgg ggaggtggag 1860  
 25 gttgcagtga gctgagatca cgcattgtt ctcgcctg ggcgcacagag caagactcca 1920  
 tttcaaaaaa aaaaaaaaaa aaaaaaaatcc actcatataa aaggtgagct cagctcaetg 1980  
 gtccatttct cagtggcttc tccatcctca ttgcacacc tcagaggat aaggcagttg 2040  
 aacctgatga gcaagaatta taacagcaag gaaacattaa tgcttagaat tctgagatcc 2100  
 30 agcacaactc agtctgtggg agctcagctc gctgcccagg gataggtatg acctatgtct 2160  
 gccttaggct gctgggagat gccattctcc agtttcagaa gcaggcaggg caaaggtaa 2220  
 gactgtggta ttggggctt ttggctctga aggatcctgg aaccactgtat ttgggtttat 2280  
 tccctccagg gtctaaagag aacaagaggt gctagctt accaaaacag atggtagaga 2340  
 35 gagttgctgg ctataaaaa agctcttca tctttttaatt cacctttct tttcacctct 2400  
 ttaaccactc ctcaggaaca gaacacttct aggactgggg gtcttttagc tccataagca 2460  
 agttagcaga tgggacaagt tagtctttc tccctagaaa caaaggggat gcccagtggt 2520  
 ttccctttgc ttcccaacct aaaatttcaa gtttaataaa atagcaatta gcagaagtga 2580  
 40 ccaaattggg agataattat cagtcatgag gaaagacaca gatttcggc ataaagaatg 2640  
 taagggttat aagttagaaac ttctataaac ctaaatgatg ttatagaatt attttgagc 2700  
 aggagcagaa agattaaata tgcacttca atacttctaa atcagaaata ggaagattaa 2760  
 aaccacagaa cagtttgtga tttctattgc tggtagctag gtatctact ctgtccactc 2820  
 45 ttgttcaagt atctaactct tctggaaacc aaataggctt tagaagagat tatttttat 2880  
 tcctatcagt ataataactaa aatgttaactt ttatcattc tggttttaa aagataaaaca 2940  
 gtttagccca tctctccaga gagcaaacat aggaatatga ctcaggagcc tcctagggt 3000  
 tatacatcagc ctcacaccc gttcccccct ccaacccaca gcctttgcctt ccaggtggca 3060  
 50 ggattactac ttgcctctt cagcagcatc tactcttaggc atattgatca ttttagacac 3120  
 tgggagaaga gaaacctaaa ctaggaggaa aagacagagc ctccacttag ttttgggagg 3180  
 ggtatggcaga cagtcaagga gatgagcgtc ctaaggcatg ttggatagg gtcagatgca 3240  
 ccacccatgg agaggttgtt caacacaaaag acatggaaagg ttagaggttt gtcaacaaaa 3300  
 55 agacatggaa gtttaggtt gtcaacacaa agacatggaa gatttagaggt ttgtcaacac 3360

aaagatacag gaagaatggg ctgcagaaga ttttagatgtt ttccatttgg gcacattta 3420  
 cttagctgga gaactagtt taaaacagcc tggtaggaa aattagaagc aagctggatg 3480  
 5 cagtggctca tgcctgtaat cccaacactt ttgggaggc caggcaggag gatcacttgg 3540  
 gcccaggagg tcaaggctgc agcgagctga gatcacacca ctgcactcca gcctgggtg 3600  
 atagaacaag accctgtctc aaaaaaaaaaaa aaaaacaaca aaaacctaga attgaggagt 3660  
 10 tgtacctcca ttggcttcct cactccaaaa taggtgctga tccttcstat tcctatttt 3720  
 tggccacccccc tgggtgttgt gtcaccagcc tgtagtttgc agtagcttgc ggcataaggct 3780  
 gccaatctg agcaaacacc agtgaggctc tattgagcaa gaccaagtcc tcaaaggcacc 3840  
 tgaaccactg tggccttctc agectacagc agtgtggctc cttacatggc cacaaaggga 3900  
 15 cacacagtga caaaaggctc ggaatgttac aatggtaaaa tgagtatct caaatccact 3960  
 gacagatata aaataggctt agagaggaaa agtgcctct ggtcaagtag atcatggcag 4020  
 catgaattcc aactcacccccc ttacgaaact ccaacttcta tgtttatctt tgttacttcc 4080  
 actttttac aacctgnccag aggcatcccc taaatcaggc ccaatatcag tattttttt 4140  
 gtgtgtgcca atttgttat cacatcccta tgaagttgaa aaataaagtt aattttgacc 4200  
 20 aaaaag 4205

<210> 2  
 <211> 4089  
 25 <212> DNA  
 <213> Unknown Organism

30 <220>  
 <223> Description of Unknown Organism:

unknown

35 <400> 2  
 cgtcacagtc ggcaggcagc atgggaaagg gagggAACCA gggcgagggg gcccggcagc 60  
 gcgagggtgct ggtgcccacc ttcagctggg aggagattca gaagcataac ctgcgcaccc 120  
 40 acaggtggct ggtcattgac cgcaagggtt acaacatcac caaatggtcc atccagcacc 180  
 cggggggcca gcgggtcatc gggcaactacg ctggagaaga tgcaacggat gccttcccg 240  
 ccttccaccc tgacctggaa ttctggca agttcttggaa acccctgtctt atggtaac 300  
 tggccccggc ggagcccagc caggaccacg gcaagaactc aaagatcaact gaggacttcc 360  
 45 gggccctgag gaagacggct gaggacatga acctgttcaa gaccaaccac gtgttcttcc 420  
 tcctccctt gggccacatc atcggccctgg agagcattgc atggttcaact gtctttact 480  
 ttggcaatgg ctggatttctt accctcatca cggcccttgc cttgttacc tctcaggccc 540  
 aagctggatg gctgcaacat gattatggcc acctgtctgt ctacagaaaa cccaaagtgg 600  
 50 accacccctgtt ccacaaattc gtcattggcc actttaaagggg tgcctctgcc aactgggtgg 660  
 atcatcgcca cttccagcac caccccaagc ctaacatctt ccacaaggat cccgatgtga 720  
 acatgctgca cgtgtttgtt ctggggcaat ggcagcccat cgagtacggc aagaagaagc 780  
 tgaaataacct gcccataat caccacacg aatacttctt cctgattggg ccggccgtgc 840  
 55 tcataccat gtatttccat taccagatca tcatgaccat gatcgccat aagaactggg 900

tggacctggc ctggggccgtc agctactaca tccgggttctt catcacctac atccctttct 960  
 acggcatcct gggagccctc ctttcctca acttcatca gttcctggag agccactgg 1020  
 5 ttgtgtgggt cacacagatg aatcacatcg tcatggagat tgaccaggag gcctaccgtg 1080  
 actggttcag tagccagctg acagccacct gcaacgtgga gcagtccctc ttcaacgact 1140  
 ggttcagtgg acaccttaac ttccagattg agcaccacct cttccccacc atgccccggc 1200  
 10 acaacttaca caagatcgcc ccgcgtggta agtctctatg tgccaagcat ggattgaat 1260  
 accaggagaa gccgctactg agggccctgc tggacatcat caggtccctg aagaagtctg 1320  
 ggaagctgtg gctggacgcc tacccatca aatgaagcca cagccccgg gacaccgtgg 1380  
 ggaagggggtg caggtgggtg gatggccaga ggaatgatgg gcttttgtc tgaggggtgt 1440  
 15 ccgagaggct ggttatgca ctgctcacgg accccatgtt ggatcttct cccttctcc 1500  
 tctccctttt ctcttacat cccccccata gcacccctgcc ctcatggac ctgcccctccc 1560  
 tcacccgtca gccatcagcc atggccctcc cagtgcctcc tagcccccctc ttccaaggag 1620  
 cagagaggtg gccaccgggg gtggctctgt cctacctcca ctctctgccc ctaaagatgg 1680  
 20 gagggagacca gcggtccatg ggtctggcct gtgagttctcc cttgcagcc tggtaactag 1740  
 gcatcacccccc cgcttgggtt cttcagatgc tcttgggggtt catagggca ggtcctagtc 1800  
 gggcaggggcc cctgaccctc ccggccctggc ttcactctcc ctgacggctg ccattggtcc 1860  
 accctttcat agagaggct gcttgttac aaagctcggg tctccctctt gcagctcggt 1920  
 25 taagtacccg aggccctctt taagatgtcc agggccccag gccccggggc acagccagcc 1980  
 caaaccttgg gcccctgaaag agtcctccac cccatcacta gagtgctctg accctggct 2040  
 ttcacggggcc ccattccacc gcctcccaa cttgagcctg tgacccctggg accaaagggg 2100  
 gagtcctctcg tctcttgc tctcagcagag gcagtggcca cgttcaggga gggccggct 2160  
 30 ggcctggagg ctcagccac cctccagctt ttcctcaggg tgtcctgagg tccaagattc 2220  
 tggagcaatc tgaccctctt ccaaaggctc tggatcagc tggcagtc cagccaatcc 2280  
 ctggccattt gggcccgagg gacgtggcc ctgcaggctg caggagggca ctggagctgg 2340  
 gaggtctcgt cccagccctc cccatctcgg ggctgtgtg tggacggcgc tgcctcaggc 2400  
 35 actctccctgt ctgaacctgc ctttactgtg tttaacctgt tgctccagga tgcattctga 2460  
 taggaggggg cggcagggtt gggccctgtg acaatctgcc tttcaccaca tggccttgcc 2520  
 tcggtgccccc tgactgtcag ggagggccag ggaggcagag cgggagggag ttcaggagg 2580  
 40 aggctgcctt gaggggctgg ggagggggta cctcatgagg accaggggtgg agctgagaag 2640  
 agagggaggt gggggcttgg ggtgtggta gctgagggga cgggcaagtg agagggggagg 2700  
 gagggaaagtc ctgggaggat cctgagctgc tggatcagtc taacccacta atcgttctt 2760  
 agattcaggg gaagggcagg caccacaac tcagaatggg ggcttcggg gagggcgcct 2820  
 45 agtcccccca gctctaagca gccaggaggg acctgcattt aagcatctgg gtggccatgg 2880  
 caatggcatg ccccccagct actgtatgcc cccgacccccc gcagaggcag aatgaaccca 2940  
 tagggagctg atcgtaatgt ttatcatgtt acttccctac ccctacattt ttgtaaataa 3000  
 aataaggaat ttatctca cttccctgtt ttcctgcacg ccaatgccag gccatggtat 3060  
 50 tgggtgatag atgaggccct tctagctgg cctggccacc aggaggggtc cccatgctt 3120  
 catctctctg tatccctcc cttccctgtg gccatccac ccgcctctcc ctgctgcctc 3180  
 taaaattcat tctggggccc ggaacttggt ggaatgacc caaaaacatt ggcccatctt 3240  
 cctccctctca gcagccgacc ccagccaaat tctaaaacag ggctgagagc cacctctca 3300  
 cagctgaccc ctacccaagg agggtggcat ggaggggctt gcagagactc ttccataacat 3360  
 55 cttccccccc cagctgttcc cccaaatgtca atctggccctc ccacccctgg gccagccagc 3420

ttccacagag cgccaggcca aacagaattc ctggcctcct tggaaaggccc tggagaaggc 3480  
 cgggagcagt ggctcacgcc tgtaatccca gcaactttggg aggctgaggc gggcagatca 3540  
 5 caaagtcaag agattgagac catcctggcc aacatggtga aaccggctc ctactaaaaaa 3600  
 tacaaaaatt aggccgggtg cggtggctca cgccctgtat cccagcactt tgggaggccg 3660  
 aggcggcag atcacgaggt caggagatca agaccatcct ggctaacacg gtgaaacccc 3720  
 10 gtctctacta aaaatacaa aaatttagctg ggcgaggtgg cgggtgcctg tagtcccage 3780  
 tacgtggag gctgaggcaa gagaatggcg tgaaccccg cggggcagag cctgcagaga 3840  
 gctgagatca caccactgta ctccagcctg ggcgacagcg agactccgtc tcaaaaaaaaa 3900  
 aaaaaaaaaa aatttagctgg gcatqgtggt qcgtgcctgc agtcccagct actcaggagg 3960  
 15 ctgagacggg agaatcgctt gaacctggga ggcagaggtt gcagtgagcc aagatcgctc 4020  
 actccagcct agcgcacagag tgagactcca tctcaaataa ataaataaaat taattaatta 4080  
 aattaaatt 4089

20 <210> 3  
 <211> 1757  
 <212> DNA  
 <213> Unknown Organism

25 <220>  
 <223> Description of Unknown Organism:

30 unknown

<400> 3  
 ggccgcggcg gcagggcggg gccggagcag cggggggcg cggaggcggc gcccgggagc 60  
 35 gctttcgct tccctcgggg tcttgctcg acctcgccca ccgcctggga tccccaggac 120  
 tcgtgcgtgc agcatggcg gcgtcgggga gcccgggaccc cgggaggggac cggcgcagcc 180  
 gggggcaccg ctgcccaccc tctgctggga gcagatccgc ggcacacgacc agccggcga 240  
 40 caagtggctg gtcatcgagc gccgcgtcta cgacatcagc cgctggcaca agcggcaccc 300  
 agggggcagc cgccctcatcg gccaccacgg cgctgaggac gccacggatg cttccgtgc 360  
 cttccatcaa gatctcaatt ttgtgcgcaa gttcctacag cccctgttga ttggagagct 420  
 ggctccggaa gaaccgcagcc aggtatggacc cctgaatgcg cagctggtcg aggacttccg 480  
 45 agccctgcac caggcagccg aggacatgaa gctgtttgat gccagttcca ctttcttgc 540  
 tttctactg gcccacatcc tggccatgga ggtgctggcc tggctcctta tctacctcct 600  
 gggctctggc tgggtgccc gtgccttggc cgccctcatac ctggccatct ctcaggctca 660  
 gtccctgggt ctgcagcatg acctggggca tgcctccatc ttcaagaagt cctgggtggaa 720  
 50 ccacgtggcc cagaagttcg tggatggggca gctaaaggcc ttctccgccc actgggtggaa 780  
 cttccgcac ttccagcacc acgcacaagcc caacatttc cacaaagacc cagacgtgac 840  
 ggtggcgccc gtcttcctcc tggggggagtc atccgtcgag tatggcaaga agaaaacgcag 900  
 atacctaccc tacaaccagc agcacctgta cttcttcctg atcggccccgc cgctgctcac 960  
 cctgggtgaac tttgaagtgg aaaatctggc gtacatgtcg gtgtgcattgc agtggggcgga 1020  
 55 tttgctctgg gcccgcagct tctatgcctcc cttcttccttta tcctacctcc ctttctacgg 1080

cgtccctggg gtgcgtctct tctttgttgc tgtcagggtc ctggaaagcc actggtaatcgt 1140  
gtggatcaca cagatgaacc acatccccaa ggagatcgac cacgagaagc accgggactg 1200  
ggtcagctct cagctggcag ccacctgcaa cgtggagccc tcactttca ccaactcggt 1260  
cagcgggcac ctcaacttcc agatcgagca ccaccccttc cccaggatgc cgagacacaa 1320  
ctacagccgg gtggcccccgc tggtcaagtgc gctgtgtgcc aagcacggcc tcagctacga 1380  
agtgaagccc ttccctcaccg cgctggtgga catcgtcagg tccctgaaga agtctgtga 1440  
catctggctg gacgccttacc tccatcagtgc aaggcaacac ccaggcgggc agagaaacggc 1500  
tcagggcacc agcaaccaag ccagccccgg cgggatcgat accccccaccc ctccacccggc 1560  
cagcctgggg gtgccttgcc tgccctctgc ctactgttgt ctccccctcg gcccccccac 1620  
atgtgtattc agcagcccta tggccttgcc tctgggcctg atgggacagg ggttagagg 1680  
aggtagagcat agcacatttt cctagagcga gaattggggg aaagctgtta tttttatatt 1740  
aaaatacatt cagatgt 1757

20 <210> 4  
21 <211> 444  
22 <212> PRT  
23 <213> Unknown Organism

<220>  
<223> Description of Unknown Organism:

*30*                          *unknown*

<400> 4  
Met Ala Pro Asp Pro Val Ala Ala Glu Thr Ala Ala Gln Gly Pro Thr

1                    5                    10                    15

20 25 30

Thr Arg Arg His Pro Gly Gly Ser Arg Val Ile Ser His Tyr Ala Gly  
50 55 60

50 Gln Asp Ala Thr Asp Pro Phe Val Ala Phe His Ile Asn Lys Gly Leu  
65 70 75 80

55 Val Lys Lys Tyr Met Asn Ser Leu Leu Ile Gly Glu Leu Ser Pro Glu  
85 90 95

Gln Pro Ser Phe Glu Pro Thr Lys Asn Lys Glu Leu Thr Asp Glu Phe  
 5 100 105 110

Arg Glu Leu Arg Ala Thr Val Glu Arg Met Gly Leu Met Lys Ala Asn  
 115 120 125

10 His Val Phe Phe Leu Leu Tyr Leu Leu His Ile Leu Leu Leu Asp Gly  
 130 135 140

15 Ala Ala Trp Leu Thr Leu Trp Val Phe Gly Thr Ser Phe Leu Pro Phe  
 145 150 155 160

20 Leu Leu Cys Ala Val Leu Leu Ser Ala Val Gln Ala Gln Ala Gly Trp  
 165 170 175

25 Leu Gln His Asp Phe Gly His Leu Ser Val Phe Ser Thr Ser Lys Trp  
 180 185 190

30 Asn His Leu Leu His His Phe Val Ile Gly His Leu Lys Gly Ala Pro  
 195 200 205

35 Ala Ser Trp Trp Asn His Met His Phe Gln His His Ala Lys Pro Asn  
 210 215 220

40 Cys Phe Arg Lys Asp Pro Asp Ile Asn Met His Pro Phe Phe Phe Ala  
 225 230 235 240

Leu Gly Lys Ile Leu Ser Val Glu Leu Gly Lys Gln Lys Lys Lys Tyr  
 245 250 255

45 Met Pro Tyr Asn His Gln His Lys Tyr Phe Phe Leu Ile Gly Pro Pro  
 260 265 270

Ala Leu Leu Pro Leu Tyr Phe Gln Trp Tyr Ile Phe Tyr Phe Val Ile  
 275 280 285

50 Gln Arg Lys Lys Trp Val Asp Leu Ala Trp Met Ile Thr Phe Tyr Val  
 290 295 300

Arg Phe Phe Leu Thr Tyr Val Pro Leu Leu Gly Leu Lys Ala Phe Leu  
 305 310 315 320

Gly Leu Phe Phe Ile Val Arg Phe Leu Glu Ser Asn Trp Phe Val Trp  
 5 325 330 335  
  
 Val Thr Gln Met Asn His Ile Pro Met His Ile Asp His Asp Arg Asn  
 10 340 345 350  
  
 Met Asp Trp Val Ser Thr Gln Leu Gln Ala Thr Cys Asn Val His Lys  
 15 355 360 365  
  
 Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu  
 20 370 375 380  
  
 His His Leu Phe Pro Thr Met Pro Arg His Asn Tyr His Lys Val Ala  
 25 385 390 395 400  
  
 Pro Leu Val Gln Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Ser  
 30 405 410 415  
  
 Lys Pro Leu Leu Ser Ala Phe Ala Asp Ile Ile His Ser Leu Lys Glu  
 35 420 425 430  
  
 Ser Gly Gln Leu Trp Leu Asp Ala Tyr Leu His Gln  
 40 435 440  
  
 <210> 5  
 <211> 444  
 <212> PRT  
 45 <213> Unknown Organism  
  
 <220>  
 <223> Description of Unknown Organism:  
  
 50 <400> 5  
 Met Gly Lys Gly Gly Asn Gln Gly Glu Gly Ala Ala Glu Arg Glu Val  
 55 1 5 10 15  
  
 Ser Val Pro Thr Phe Ser Trp Glu Glu Ile Gln Lys His Asn Leu Arg  
 60 20 25 30

	Thr Asp Arg Trp Leu Val Ile Asp Arg Lys Val Tyr Asn Ile Thr Lys		
5	35	40	45
	Trp Ser Ile Gln His Pro Gly Gly Gln Arg Val Ile Gly His Tyr Ala		
	50	55	60
10	Gly Glu Asp Ala Thr Asp Ala Phe Arg Ala Phe His Pro Asp Leu Glu		
	65	70	75
	Phe Val Gly Lys Phe Leu Lys Pro Leu Leu Ile Gly Glu Leu Ala Pro		
15	85	90	95
20	Glu Glu Pro Ser Gln Asp His Gly Lys Asn Ser Lys Ile Thr Glu Asp		
	100	105	110
	Phe Arg Ala Leu Arg Lys Thr Ala Glu Asp Met Asn Leu Phe Lys Thr		
25	115	120	125
	Asn His Val Phe Phe Leu Leu Leu Ala His Ile Ile Ala Leu Glu		
	130	135	140
30	Ser Ile Ala Trp Phe Thr Val Phe Tyr Phe Gly Asn Gly Trp Ile Pro		
	145	150	155
	160		
35	Thr Leu Ile Thr Ala Phe Val Leu Ala Thr Ser Gln Ala Gln Ala Gly		
	165	170	175
40	Trp Leu Gln His Asp Tyr Gly His Leu Ser Val Tyr Arg Lys Pro Lys		
	180	185	190
	Trp Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala		
	195	200	205
45	Ser Ala Asn Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro		
	210	215	220
50	Asn Ile Phe His Lys Asp Pro Asp Val Asn Met Leu His Val Phe Val		
	225	230	235
	240		
55	Leu Gly Glu Trp Gln Pro Ile Glu Tyr Gly Lys Lys Lys Leu Lys Tyr		
	245	250	255

5	Leu Pro Tyr Asn His Gln His Glu Tyr Phe Phe Leu Ile Gly Pro Pro			
	260	265	270	
10	Leu Leu Ile Pro Met Tyr Phe Gln Tyr Gln Ile Ile Met Thr Met Ile			
	275	280	285	
15	Val His Lys Asn Trp Val Asp Leu Ala Trp Ala Val Ser Tyr Tyr Ile			
	290	295	300	
20	Arg Phe Phe Ile Thr Tyr Ile Pro Phe Tyr Gly Ile Leu Gly Ala Leu			
	305	310	315	320
25	Leu Phe Leu Asn Phe Ile Arg Phe Leu Glu Ser His Trp Phe Val Trp			
	325	330	335	
30	Val Thr Gln Met Asn His Ile Val Met Glu Ile Asp Gln Glu Ala Tyr			
	340	345	350	
35	Arg Asp Trp Phe Ser Ser Gln Leu Thr Ala Thr Cys Asn Val Glu Gln			
	355	360	365	
40	Ser Phe Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu			
	370	375	380	
45	His His Leu Phe Pro Thr Met Pro Arg His Asn Leu His Lys Ile Ala			
	385	390	395	400
50	Pro Leu Val Lys Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Glu			
	405	410	415	
55	Lys Pro Leu Leu Arg Ala Leu Leu Asp Ile Ile Arg Ser Leu Lys Lys			
	420	425	430	
	Ser Gly Lys Leu Trp Leu Asp Ala Tyr Leu His Lys			
	435	440		

5           <210> 6  
          <211> 445  
          <212> PRT  
          <213> Unknown Organism

10          <220>  
          <223> Description of Unknown Organism:  
  
               unknown

15          <400> 6  
          Met Gly Gly Val Gly Glu Pro Gly Pro Arg Glu Gly Pro Ala Gln Pro  
               1                   5                   10                   15

20          Gly Ala Pro Leu Pro Thr Phe Cys Trp Glu Gln Ile Arg Ala His Asp  
               20                   25                   30

25          Gln Pro Gly Asp Lys Trp Leu Val Ile Glu Arg Arg Val Tyr Asp Ile  
               35                   40                   45

30          Ser Arg Trp Ala Gln Arg His Pro Gly Gly Ser Arg Leu Ile Gly His  
               50                   55                   60

35          His Gly Ala Glu Asp Ala Thr Asp Ala Phe Arg Ala Phe His Gln Asp  
               65                   70                   75                   80

40          Leu Asn Phe Val Arg Lys Phe Leu Gln Pro Leu Leu Ile Gly Glu Leu  
               85                   90                   95

45          Ala Pro Glu Glu Pro Ser Gln Asp Gly Pro Leu Asn Ala Gln Leu Val  
               100                   105                   110

50          Glu Asp Phe Arg Ala Leu His Gln Ala Ala Glu Asp Met Lys Leu Phe  
               115                   120                   125

55          Asp Ala Ser Pro Thr Phe Phe Ala Phe Leu Leu Gly His Ile Leu Ala  
               130                   135                   140

          Met Glu Val Leu Ala Trp Leu Leu Ile Tyr Leu Leu Gly Pro Gly Trp  
               145                   150                   155                   160

## EP 1035 207 A1

	Val Pro Ser Ala Leu Ala Ala Phe Ile Leu Ala Ile Ser Gln Ala Gln		
	165	170	175
5	Ser Trp Cys Leu Gln His Asp Leu Gly His Ala Ser Ile Phe Lys Lys		
	180	185	190
10	Ser Trp Trp Asn His Val Ala Gln Lys Phe Val Met Gly Gln Leu Lys		
	195	200	205
15	Gly Phe Ser Ala His Trp Trp Asn Phe Arg His Phe Gln His His Ala		
	210	215	220
20	Lys Pro Asn Ile Phe His Lys Asp Pro Asp Val Thr Val Ala Pro Val		
	225	230	235
	Phe Leu Leu Gly Glu Ser Ser Val Glu Tyr Gly Lys Lys Lys Arg Arg		
	245	250	255
25	Tyr Leu Pro Tyr Asn Gln Gln His Leu Tyr Phe Phe Leu Ile Gly Pro		
	260	265	270
30	Pro Leu Leu Thr Leu Val Asn Phe Glu Val Glu Asn Leu Ala Tyr Met		
	275	280	285
35	Leu Val Cys Met Gln Trp Ala Asp Leu Leu Trp Ala Ala Ser Phe Tyr		
	290	295	300
40	Ala Arg Phe Phe Leu Ser Tyr Leu Pro Phe Tyr Gly Val Pro Gly Val		
	305	310	315
	Leu Leu Phe Phe Val Ala Val Arg Val Leu Glu Ser His Trp Phe Val		
	325	330	335
45	Trp Ile Thr Gln Met Asn His Ile Pro Lys Glu Ile Gly His Glu Lys		
	340	345	350
50	His Arg Asp Trp Val Ser Ser Gln Leu Ala Ala Thr Cys Asn Val Glu		
	355	360	365
55	Pro Ser Leu Phe Thr Asn Trp Phe Ser Gly His Leu Asn Phe Gln Ile		
	370	375	380

EP 1035 207 A1

Glu His His Leu Phe Pro Arg Met Pro Arg His Asn Tyr Ser Arg Val  
385 390 395 400

5 Ala Prc Leu Val Lys Ser Leu Cys Ala Lys His Gly Leu Ser Tyr Glu  
405 410 415

10 Val Lys Pro Phe Leu Thr Ala Leu Val Asp Ile Val Arg Ser Leu Lys  
420 425 430

15 Lys Ser Gly Asp Ile Trp Leu Asp Ala Tyr Leu His Gln  
435 440 445

20 <210> 7

<211> 18

<212> DNA

<213> Unknown Organism

25 <220>

<223> Description of Unknown Organism:

30 unknown

<400> 7

cgcctgacag cccctgct

18

35 <210> 8

<211> 20

<212> DNA

40 <213> Unknown Organism

<220>

<223> Description of Unknown Organism:

45 unknown

<400> 8

50 caggtggcca atcacaaaaat

20

<210> 9

<211> 22

55 <212> DNA

5           <213> Unknown Organism

10          <220>

15          <223> Description of Unknown Organism:

20                 unknown

25          <400> 9

25          ctcaaaagtgg aaccatctgc ta

30          <210> 10

30          <211> 20

35          <212> DNA

35          <213> Unknown Organism

40          <220>

40          <223> Description of Unknown Organism:

45                 unknown

50          <400> 10

50          ggaaacccag tccatgttcc

55          <210> 11

55          <211> 21

55          <212> DNA

55          <213> Unknown Organism

60          <220>

<223> Description of Unknown Organism:

5                   unknown

<400> 12

ctcaagctcc cctctgcct

19

10                  

<210> 13

<211> 19

<212> DNA

15                  <213> Unknown Organism

<220>

<223> Description of Unknown Organism:

20                  unknown

<400> 13

tcagaagcat aacctgcgc

19

25                  

<210> 14

<211> 20

<212> DNA

<213> Unknown Organism

35                  <220>

<223> Description of Unknown Organism:

40                  unknown

40                  

<400> 14

ccagttcacc aatcagcagg

20

45                  <210> 15

<211> 20

<212> DNA

50                  <213> Unknown Organism

<220>

<223> Description of Unknown Organism:

55                  unknown

5 <400> 15  
ccccatgcga ttggtaact 20

10 <210> 16  
<211> 20  
<212> DNA  
<213> Unknown Organism

15 <220>  
<223> Description of Unknown Organism:  
unknown

20 <400> 16  
tc:agggcag gtatttca 20  
gc

25 <210> 17  
<211> 19  
<212> DNA  
<213> Unknown Organism

30 <220>  
<223> Description of Unknown Organism:  
unknown

35 <400> 17  
agcccatcgaa gtacogcaa 19  
gt

40 <210> 18  
<211> 20  
<212> DNA  
<213> Unknown Organism

45 <220>  
<223> Description of Unknown Organism:  
unknown

50 <220>  
<223> Description of Unknown Organism:  
unknown

55 <400> 18  
cc:cagaaca aaaggccatc 20

5 <210> 19  
<211> 18  
<212> DNA  
<213> Unknown Organism

10 <220>  
<223> Description of Unknown Organism:  
  
15 unknown

15 <400> 19  
tcttgctcgg acctcgcc 18

20 <210> 20  
<211> 20  
<212> DNA  
25 <213> Unknown Organism

25 <220>  
<223> Description of Unknown Organism:  
  
30 unknown

30 <400> 20  
gtgatccaca cgaaccagtg 20

35 <210> 21  
<211> 19  
<212> DNA  
<213> Unknown Organism

40 <220>  
<223> Description of Unknown Organism:  
  
45 unknown

45 <400> 21  
gaagaaccca gccaggatg 19

50 <210> 22  
<211> 22

5           <211> 20  
           <212> DNA  
           <213> Unknown Organism

10           <220>  
           10    <223> Description of Unknown Organism:

unknown

15           <400> 22  
           acagctttcc cccaaattctc

20

20

### Claims

25           1. An isolated cDNA molecule selected from the group consisting of

- 25           (a) a polynucleotide having at least a 65 % homology, preferably at least a 80 % homology with a polynucleotide  
           encoding a polypeptide selected from the group consisting of the polypeptides of SEQ ID NO: 4, SEQ ID NO:  
           5 and SEQ ID NO: 6;  
           (b) a polynucleotide having at least a 65 % homology, preferably at least a 80 % homology with a polynucleotide  
           which by virtue of the redundancy of the genetic code, encodes the same polypeptide selected from the group  
           consisting of the polypeptides of SEQ ID NO: 4, SEQ ID NO: 5 and SEQ ID NO: 6;  
           (c) a DNA molecule capable of hybridization under stringent conditions to a DNA molecule according to (a) or  
           (b);  
           (d) a polynucleotide which is complementary to the polynucleotide of (a), (b) or (c); and  
           (e) a oligonucleotide comprising at least 15 consecutive nucleotides of the polynucleotide of (a), (b), (c) or (d).

30           2. An isolated cDNA molecule selected from the group consisting of

- 30           (a) a polynucleotide having at least a 65 % homology, preferably at least a 80 % homology with a polynucleotide  
           sequence selected from the group consisting of the polynucleotides of SEQ ID NO: 1, SEQ ID NO: 2 and SEQ  
           ID NO: 3;  
           (b) a DNA molecule capable of hybridization under stringent conditions to a DNA molecule according to (a);  
           (c) a polynucleotide which is complementary to the polynucleotide of (a) or (b);  
           (d) a oligonucleotide comprising at least 15 consecutive nucleotides of the polynucleotide of (a), (b) or (c); and  
           (e) a DNA which is synonymous to the DNAs of (a), (b), (c) or (d) due to the degeneracy of the genetic code.

35           3. A DNA comprising a nucleotide sequence with at least a 65 % homology with the nucleotide sequences as defined  
           in claim 1 or 2.

40           4. A recombinant vector comprising the DNA as claimed in any of claims 1 to 3.

45           5. A transgenic host cell comprising the DNA as claimed in any of claims 1 to 3.

50           6. A transgenic host cell transformed by the DNA according to any of claims 1 to 3 or the vector according to claim  
           4, a corresponding transgenic organism or a corresponding transgenic knock-in or knock-out animal model.

55           7. A polypeptide comprising at least 65 % of a polypeptide sequence selected from the group consisting of the  
           polypeptides of SEQ ID NO: 1, SEQ ID NO: 2 and SEQ ID NO: 3, or its salt.

8. A polypeptide comprising a polypeptide sequence with at least a 85 % homology with the polypeptide sequence as claimed in claim 7, or its salt.
9. A peptide comprising at least 15 consecutive amino acids of the polypeptide as claimed in claim 7, or its salt.
- 5 10. A polypeptide having substantially the same amino acid sequence as the polypeptide as claimed in claim 7, or having a variant of the amino acid sequence of the polypeptide as claimed in claim 7 with a deletion, addition or substitution of 1 to 10 amino acids, or its salt.
- 10 11. A process for producing a polypeptide comprising expressing from the host cell of claim 5 or 6 a polypeptide encoded by the DNA as claimed in any of claims 1 to 3.
12. An antibody against the polypeptide of any of claims 7 to 10.
- 15 13. A oligonucleotide primer having a nucleotide sequence selected from the group consisting of the nucleotide sequences of SEQ ID NO: 7 to SEQ ID NO: 22.
14. A method of screening for modulators in known assays using constructs or of screening for interacting proteins or factors using state of the art technologies.
- 20 15. A method of screening chemical libraries comprising transformed cell lines.
16. A compound which alters or reacts with at least one epitope of the proteins and which is obtained by screening methods as claimed in claim 14 or 15.
- 25 17. The use of the antibodies according to claim 12 for diagnostic or therapeutic purposes.
18. A pharmaceutical composition comprising as an effective component an effective amount of the peptide as claimed in any of claims 7 to 10, or its salt, and a pharmaceutically acceptable carrier or diluent.

30

35

40

45

50

55

Fig. 1

**FADS1** 1 .MGPDPVAAETAAQGPTP.RYFTWDEVAORSGCERWLVIDRKVYNIEETRHPGGSR  
**FADS2** 1 MGRGGNQG..ECAAKREVSLPTFSEWEIOKHNLRTDRWLVIDRKVYNITWNSICHPGGCR  
**FADS3** 1 MGGVGEPGPREGPAOPGAP.PTFCWECIRAHDOQPGDKWLVIERRVYDIESRWORHPGGSR  
**Borago** 1 .....MAQINKYVTSDELINHDKPGDWISIOGKAYDVSOWADHPGGSF  
**Helianthus** 1 ....MVSPSIEVLNSIDGKTYTSKELAKHNPNNDLWISILGKVNTIEWAEHPGGDA  
**Cytochrome b5** 1 .....MAEDSDEAKYITDEEIKHNEKSATWLLIEHKVYDLTRLLEHPGGLE

**FADS1** 58 VISHYAGQDATDPFVAFHINKGLVKKYNNSLLIGELSPLOPSPPPTNKGELDIFRCLRA  
**FADS2** 59 VIGHYAGEDATDAFRAFHPOLEEVCKSLPLLIGELAPEEPSODHGINSKLLDFRFLRK  
**FADS3** 61 EIGHEGAEDATDAFRAFHDLNEVRKELPFLLLIGELAPEEPSOQGPLNACLVEDFRALHQ  
**Borago** 47 PLKSIAGQEVTDADFVAFHPAST...KNDKQFTTCYLNKYSV.....SEVSKDYRQJUT  
**Helianthus** 57 PLINLAGQDVTDADFIAFHPGTAV...KRLQDQTCYHLKDYQV.....SDISKDYRQIAS  
**Cytochrome b5** 50 VLPREOAGGEDATENFEDVGESTDA.EMSKTFIIGELHPDDR.....PAENKPPPETLIT

**FADS1** 118 TVERMGLKIANHVPFLAYLHILLLDGAWLTHVFGTSIPELLCAVLLSAVQAQAGWL  
**FADS2** 119 TAEDMNLKTNHVPFLLAHILAHLESLAWEVFGNGWPTLITAFVLTSAQAGWL  
**FADS3** 121 AAEDMNLKASPTFEAFLGHILALMVLAWLILYLGPGWVPSALAFLAISQAQSWCL  
**Borago** 99 EFSIAGLYDIKQHIMFATCFIAMLFA.SVTCV..LECEGVIVHESGCCUGFLWIOSGWV  
**Helianthus** 109 EFAVAGMEIAGKGEVYISLCFVSLLLSACVLYV.LYSGSEWIBLSCAIGLAMQHAYL  
**Cytochrome b5** 102 TESSSSSTWNVIPAISSAVAVALMYRPTAED.....

**FADS1** 178 QHDAGHLSVSVTSKAWNHEKHFVIGPLKGAPASWWNEHFOHHAKPNCEPKDPDLMHPT  
**FADS2** 179 QHDAGHLSVMRKPAKWNHLEKHFVIGHLKGSASAWWNEHFOHHAKPNIFEKDPDVHMLHV  
**FADS3** 181 QHDAGHLSVTSKGSWNHVAQKFVUMGOLKGPSAEWWNEHFOHHAKPNIFHKDPDVHMLPV  
**Borago** 158 EHDAGHMLVVSDSRNLKFLYGFIAANCLGICISICWWVNAHNHHIACNSLEYDPDLYIPE  
**Helianthus** 168 QHDAGHMQ.MATRGWVAGLGFILGNCCAGISIANWWVNAHNHHIACNSLDYDPDLOBLPL

**FADS1** 238 FFA....LGQISVELGKO....KLKYMPYNEQHQIYFFLIGPPALLPLYFQIYIFYEV.  
**FADS2** 239 FV....LGEVQEKEYGKE....KLKYLPPYNEQHQIYFFLIGPPPLIPLYFQIQLITL.  
**FADS3** 241 FV....LGQISSEVEYGKA....KRRYLPPYNEQHQIYFFLIGPPPLIPLYFQIQLITL  
**Borago** 218 LVSSSKFFGSLTSEFYKELTFDSLRSRFFVSYQHNTFPIPMCMARQNWVOSLIMLTKR  
**Helianthus** 228 LVASSKLFNSITSVSYGROLTDFPLARFFVSYQHLYTPIPMCVAVNLYQTILWISKR

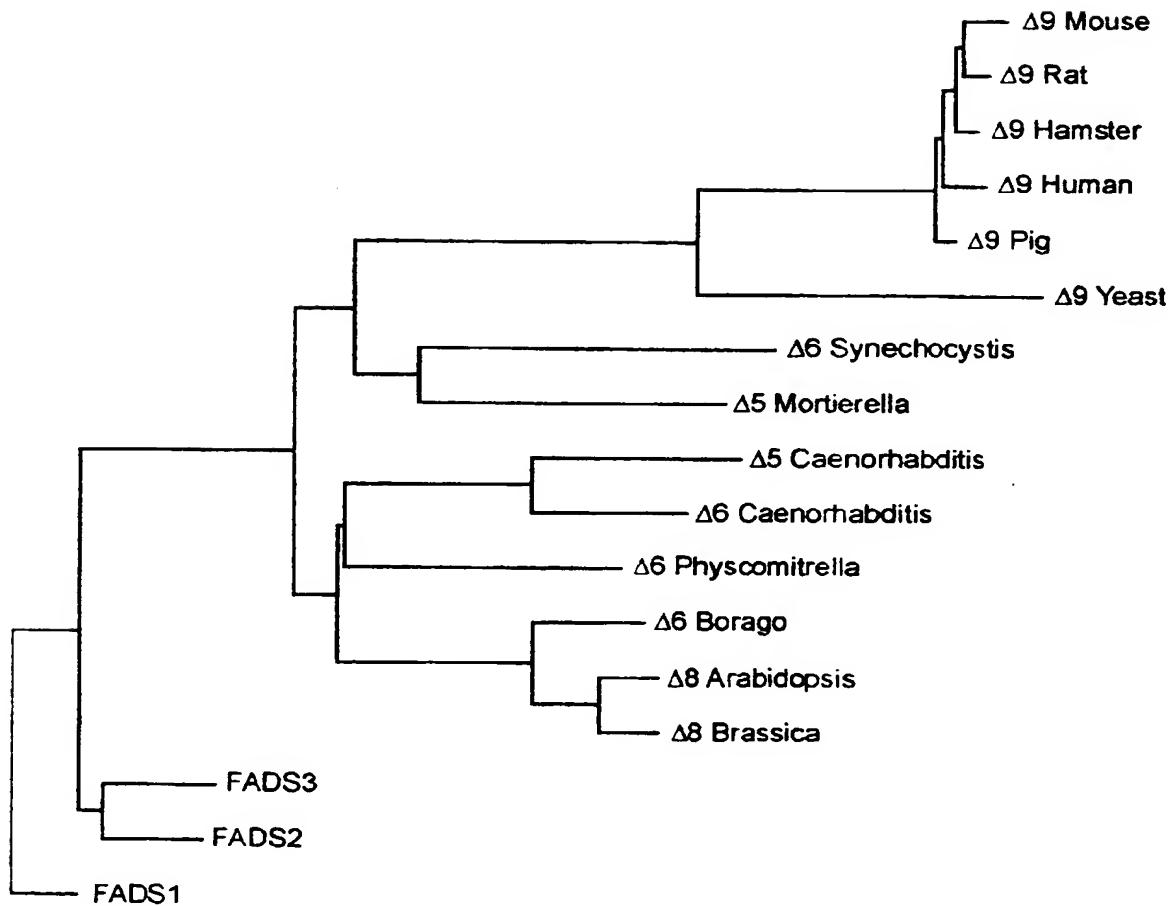
**FADS1** 289 .ORKAVWDLAWVTFYVREFELTYVPLIGKAKFLGLEFIVRQLES.WFWVVWTQMNHIP..  
**FADS2** 289 .V.EONWVDLAWVSYIYREFELTYIPTFYGILCALLFLNPIRQLES.WFWVVWTQMNHIP..  
**FADS3** 290 .VCNOQADLWVPASEYVREFELSYLPTFYGIVPCVLLFEGAVRQLES.WFWVVWTQMNHIP..  
**Borago** 278 NVSYRABELLGOVVESTWVPLLVSCLPNNGERIMFVIAISLVTGNOOVOFS.LNHESSSV  
**Helianthus** 288 KIPDREGDNILGDTIETWPLLVSRIPNMPERVAFVUSPCUTGIQHQLQFT.LNHPSGDV

**FADS1** 346 HIDEDERNMDWSTQLOCATCNVHKS.FNDWFSGHLENFQIEHHLFPMMPRHNLKIAPLVKS  
**FADS2** 346 EIDQAYRDWESSQLTATCNVQSEENWFWSGCHLENFQIEHHLFPMMPRHNLKIAPLVKS  
**FADS3** 347 EIGHEKERDWESELQATCNVESPSETWFWSGCHLENFQIEHHLFPMMPRHNLKIAPLVKS  
**Borago** 337 YVGPKGNWFEAQDGLDIDSCPPWADWFEGELQFQIEHHLFPMMPQNLKISPIVIL  
**Helianthus** 347 YVGPPKGNWPERQTRGIDIDACSSADWFEGELQFQIEHHLFPRLPRCHLRSISPICRE

**FADS1** 406 LCAKHGIEYQSKPLLSAFADITESLK-SGOLWLDAYLH.....  
**FADS2** 406 LCAKHGIEYQSKPLLSAFADITESLK-SGOLWLDAYLH.....  
**FADS3** 407 LCAKHGLSYLVKPFLALVDIVRSLKNSGDIWLDAYLH.....  
**Borago** 397 LQKHNLPYKASFSKANMTLRTLRTMELQARDITKPLPKNLVWEALEHTHG  
**Helianthus** 407 LQKHNLPYVSLSEYDANVTLVLTETALQARDITNPAPQNLAEEFNTHG

□ transmembrane region  
 □ conserved histidine box  
 ▼ invariant amino acid residue

**Fig.2**



**Fig. 3****FADS1 cDNA**

CACTCCTGGAGCCCCGGACCCCGAGCACGCGCCTGACAGCCCCCTGCTGGCCCGCGCGCGCG  
 TCGCCAGGCCAGCTATGCCCGGCGACCGGGTGGCCGAGACCGCGGCTCAGGGACCTACCC  
 GCGCTACTTCACCTGGGACGAGGTGGCCAGCGCTCAGGGTGCAGGGAGCGGTGGCTAGTGATC  
 GACCGTAAGGTGTACAACATCAGCGAGTTCAACCGCCGGCATCCAGGGGCTCCC GGTCATCA  
 GCCACTACGCCGGGAGGATGCCACGGATCCCTTG1GGCCTTCCACATCAACAAGGGCTTGT  
 GAAGAAGTATATGAACTCTCTCCTGATTGGAGAACTGTCTCCAGAGCAGGCCAGCTTGAGGCC  
 ACCAAGAATAAAAGAGCTGACAGATGAGTCCGGGAGCTGCGGGCACAGTGGAGCGGATGGGGC  
 TCATGAAGGCCAACCATGTCTTCTCCTGCTGTACCTGCTGCACATCTTGCTGCTGGATGGTGC  
 AGCCTGGCTACCCTTGGGCTTGGGACGTCCTTTGCCCTCCTCTGTGCGGTGCTG  
 CTCAGTGCAGTCAGGCCAGGCTGGCTGGCTGCAGCATGACTTGGCACCTGTCGGTCTTCA  
 GCACCTCAAAGTGGAACCATCTGCTACATCATTGTGATTGGCACCTGAAGGGGCCCGC  
 CAGTTGGTGGAACCATGCACTTCCAGCACCATGCCAGCCAAGCCAACTGCTCCGCAAAGACCC  
 GACATCAACATGCATCCCTTGCCTGGGGAGATCCTCTGTGGAGCTTGGAAAC  
 AGAAGAAAAATATATGCCGTACAACCACAGCACAAATACTTCTCTAATTGGGCCCGC  
 CTTGCTGCCTCTACTTCCAGTGGTATATTCTATTGTATCCAGCGAAAGAAGTGGGTG  
 GACTTGGCCTGGATGATTACCTCTACGTCCGCTTCCCTCACTTATGTGCCACTATTGGGGC  
 TGAAAGCCTCTGGGCCCTTCTCATAGTCAGGTTCTGGAAAGCAACTGGTTGTGGT  
 GACACAGATGAACCATATTCCCATGACATTGATCATGACCGAACATGGACTGGTTCCACC  
 CAGCTCCAGGCCACATGCAATGTCACAAGTCTGCCTCAATGACTGGTTCAGTGGACACCTCA  
 ACTTCCAGATTGAGCACCATTTCACGATGCCCTGACACAATTACCAAAAGTGGCTCC  
 CCTGGTCAGTCCTGTGCTGCCAGCATGGCATAGAGTACAGTCCAAGGCCCTGCTGTCAGCC  
 TTCGCCGACATCATCCACTCAAAAGGAGTCAGGGCAGCTGGCTAGATGCCATCTTCACC  
 AATAACAACAGCCACCCCTGCCAGTCTGGAGAAAGAGGAGGAAGACTCTGGAGCCAAGGCAGAG  
 GGGAGCTTGAGGGACAATGCCACTATAGTTAATACTCAGAGGGGTTGGGTTGGGACATAA  
 AGCCTCTGACTCAAACCTCCCTTTATCTCTAGCCACAGTTCTAAGACCCAAAGTGGGGG  
 TGGACACAGAACGTCCTAGGAGGGAGGAGCTGTGGGCAAGGGGTGTAATTATTCCTTTT  
 CTAGTTGGCACATGCAGGTAGTTGGTGAACAGAGAACAGGAGGGTAACAGAAGAGGAGGG  
 ACCTACTGAACCCAGAGTCAGGAAGAGATTAAACACTAAAATCCACTCATGCCGGCGTGGTG  
 CACGCCCTGTAATCCCAGCTACCCAGGAGGCTGAGGCAGGAGAACGCTGAACCGGGGAGGT  
 GGAGGTTGCAGTGAGATCACGCCATTGTACTCCAGCCTGGGCAAGAGAACACTCCA  
 TTTCAAAAAAAAAAAATCCACTCATAAAAGGTGAGCTCAGCTACTGGTC  
 ATTTCTCAGTGGCTTCTCCATCTCATTGCAAACCTCAGAGGGATAAGGCAGTTGAACCTGAT  
 GAGCAAGAATTATAACAGCAAGGAAACATTAAATGCTTAGAATTCTGAGATCCAGCACAACTCAG  
 TCTGTGGGAGCTCAGTCGCTGCCAGGGATAGGTATGACCTATGTCTGCCTTAGGCTGCTGG  
 AGATGCCATTCTCCAGTTCAAGCAGGCAGGGCAAGGTCAAGACTGTGGTATTGGGTCTT  
 TTGGCTCTGAAGGATCCTGGAACCACTGATTGGTTATTCCCTCCAGGGCTAAAGAGAAC  
 AGAGGTGCTAGCTTACCAAAACAGATGGTAGAGAGAGTTGCTGGCTATTAAAAGCTTT  
 CATCTTTAATTCAACCTCTTCAACCTCTTAAACACTCCTCAGGAACAGAACACTCTAG  
 GACTGGGGGTCTTCTAGCTCCATAAGCAAGTGAGCAGATGGGACAAGTTAGTCTTTCTCC  
 GAAACAAAGGGGATGCCAGTGGTTCCCTTGCTTCCAAACCTAAAATTCAAGTTAATAAA  
 ATAGCAATTAGCAGAACAGTGAACAAATTGGGAGATAATTACAGTCATGAGGAAAGACACAGATT  
 TCGGTCAAAAGAATGTAAGGGCTATAAGTAGAAACTTCTATAACCTAAATGATGTTAGAA  
 TTATTTGAGCAGGAGCAGAAAGATTAAATATGATCACTTCATACTCTAAATCAGAAATAGG  
 AAGATTAAACACAGAACAGTTGTGATTCTATTGCTGGTAGCTAGGTATCTTACTCTGTCC  
 ACTCTTGTCAAGTATCTAACTCTCTGGAAACCAAATAGGCTTTAGAAGAGATTATCCTATAT  
 TCCTATCAGTATAACTAAATGTAACCTTTAATCATCTGGTTTAAAGATAAACAGTT

**Fig. 3 cont.**

AGCCCATCTCTCCAGAGAGCAAACATAGGAATATGACTCAGGAGCCTCCTAGGGCTTATCATCA  
 GCCCTCACACCCGTTCCCCCTCCAACCCACAGCCTTGCCTTCAGGTGGCAGGATTACTACTT  
 TGCCTCTTCAGCAGCATCTACTCTAGGCATATTGATCATTAGACACTGGGAGAAGAGAACCT  
 CAAACTAGGAGGAAAAGACAGAGCCTCCACTTAGTTGGGAGGGGATGGCAGACAGTCAGGA  
 GATGAGCGTCTTAAGGCATGTTGGGATAGGGTCAGATGCACCACCCATGGAGAGGTTGTCAAC  
 ACAAAAGACATGGAAGGTTAGAGGTTGTCAACAAAAAGACATGGAAGGTTAGGTTGTCAACAC  
 AAAGACATGGAAGGATTAGAGGTTGTCAACACAAAGATACTGGAAAGAATGGGCTGCAGAAGATT  
 TAGATGTTTCCATTGGCACATTTACTTAGCTGGAGAACTAGGTTAAACAGCCTGGGTA  
 GGAAAATTAGAACAGCTGGATGCAGTGGCTCATGCCGTAACTCCAAACACTTTGGGAGGTC  
 CAGGCAGGAGGATCACTGGGCCAGGAGGTCAAGCCTGCAGCGAGCTGAGATCACACCACTGC  
 ACTCCAGCCTGGGTGATAGAACAAAGACCCGTCTAAAAAAGAAAAAAACAAACAAAAACTA  
 GAATTGAGGAGTTGTACCTCCATTGGCTTCCTCACTCCAAATAGGTGCTGATCCTCCTATTC  
 CTATTCTTGCCACCTTTGGGTGTGGTGTCAACAGCCTGTTAGCCAAGTAGCTTTGGGCATA  
 GGCTGCCAATCTGAGCAAACACCAGTGAGGCTTATTGAGCAAGACCAAGTCCCAAAGCACC  
 TGAACCACTGTGGCTTCTCAGCCTACAGCAGTGTGGCTCTTACATGGCCACAAAGGGACACA  
 CAGTGACAAAAGGCTCGAATGTTACAATGGTAAAATGAGTGATCTCAAATCCACTGACAGATA  
 TAAAATAGGCTTAGAGAGGAAAAGCTGCCCTGGTCAAGTAGATCATGGCAGCATGAATTCAA  
 CTCACTTTTACGAACCTCAACTTCTATGTTATCTTGTACTTTCACTTTTACAACCTG  
 NCAGAGGCATTTTAAATCAGGCCAATATCAGTATTCTTTGTGTGCCAATTTGTTAT  
 CACATCCCTATGAAGTTGAAAAATAAGTTAATTGACCHAAAG

**Fig. 4****FADS2 cDNA**

CGTCACAGTCGGCAGGCAGCATGGGGAGGGAGGGACCAGGGCGAGGGGCCGAGCGCGA  
 GGTGTCGGTGCCCACCTCAGCTGGAGGAGATTCAAAGCATAACCTGCGCACCGACAGGTGG  
 CTGGTCATTGACCGCAAGGTTACAACATCACAAATGGTCATCCAGCACCCGGGGCCAGC  
 GGGTCATCGGGCACTACGCTGGAGAAGATGCAACGGATGCCTTCCGCGCCTTCCACCTGACCT  
 GGAATTCTGGGGCAAGTCTTGAAACCCCTGCTGATTGGTAACCTGGCCCCGGAGGAGCCCAGC  
 CAGGACACGGCAAGAACACTAAAGATCACTGAGGACTCCGGGCCCTGAGGAAGACGGCTGAGG  
 ACATGAACCTGTTCAAGACCAACCACGTGTTCTCCTCCTGGCCACATCATGCCCT  
 GGAGAGCATTGCATGGTCACTGTCTTACTTGGCAATGGCTGGATTCCATACCCATCACG  
 GCCTTGTCTTGCACCTCTCAGGCCAAGCTGGATGGCTGCAACATGATTATGGCACCTGT  
 CTGTCACAGAAAACCAAGTGGAAACCACCTGTCCACAAATTGTCATTGCCACTAAAGGG  
 TGCCTCTGCCAACCTGGTGGAAATCATGCCACTTCCAGCACGCCAACGCTAACATCTTCCAC  
 AAGGATCCCAGATGTGAACATGCTGCACGTGTTGTTCTGGCGAATGGCAGGCCATCGAGTACG  
 GCAAGAAGAAGCTGAAATACCTGCCCTACAAATCACCAGCACGAATACTCTTCTGATTGGGCC  
 GCCGCTGCTCATCCCCATGTATTCCAGTACCATCATGACCATGATGTCATAAGAAC  
 TGGGTGGACCTGGCCTGGGCCGTCACTACATCCGGTTCTCATCACCTACATCCCTTCT  
 ACGGCATCCTGGGAGCCCTCTTCTCAACTCATCAGGTTCTGGAGAGGCCACTGGTTGT  
 GTGGGTACACAGATGAATCACATGTCATGGAGATTGACCAAGGAGGCCTACCGTGA  
 AGTAGCCAGCTGACAGCCACCTGCAACGTGGAGCAGTCCTCTCAACGACTGGTCAGTGGAC  
 ACCTTAACCTCCAGATTGAGCACCCACTCTCCCCACCATGCCCGCACAACTTACACAAGAT  
 CGCCCCGCTGGTGAAGTCTCATGTGCCAACATGGCATTGAATACCAAGGAGAAGCCGCTACTG  
 AAGGGCCCTGCTGGACATCATCAGGTCCCTGAAAGAAGTCTGGGAAGCTGTGGCTGGACGCCCTACC  
 TTCACAAATGAAGCCACAGCCCCGGGACACCGTGGGAAGGGGTGCAGGTGGGTGATGGCCA  
 GAGGAATGATGGGCTTGTCTGAGGGGTGTCGAGAGGCTGGTGTATGCACTGCTCACGGAC  
 CCCATGTTGGATTTCTCCCTTCTCCTCTCTTCTTCTTCTTCAACATCTCCCCATAGCACCC  
 TGCCCTCATGGGACCTGCCCTCCCTCAGCGTCAGCCATCAGCCATGCCCTCCAGTGCCTCC  
 TAGCCCCCTTCTCCAAGGAGCAGAGAGGTGCCACCGGGGTGGCTCTGTCTACCTCCACTCT  
 CTGCCCCCTAAAGATGGGAGGAGACCAGCGGCCATGGGTCTGGCTGTGAGTCTCCCTTGAG  
 CCTGGTCACTAGGCATCACCCCCCTTGGTCTTCAGATGCTTGGGTTCATAGGGCAGG  
 TCCTAGTCGGCAGGGCCCTGACCCCTCCGGCTGGCTTCACTCTCCCTGACGGCTGCCATTG  
 GTCCACCCCTTCATAGAGAGGCCTGTTACAAAGCTGGGTCTCCCTCTGCAAGCTCGGT  
 TAAGTACCCGAGGCCCTCTTAAGATGTCAGGGCCCCAGGCCGGCACAGCCAGCCAAA  
 CCTTGGGCCCTGGAAGAGTCCTCCACCCCATCACTAGAGTGTCTGACCCCTGGCTTCACGGG  
 CCCCATTCCACCGCCTCCCCACTTGAGCCTGTGACCTTGGGACCAAAGGGGAGTCCCTCGTC  
 TCTTGTGACTCAGCAGAGGCAGTGGCACGTTCAAGGGAGGGCTGGCCTGGAGGCTCAGC  
 CCACCCCTCCAGCTTCTCAGGGTGTCTGAGGTCAAAGATTCTGGAGCAATCTGACCCCTCT  
 CCAAAGGCTCTGTTATCAGCTGGCAGTGCAGCCAATCCCTGCCATTGGCCCCAGGGGACG  
 TGGGCCCTGCAAGGCTGCAGGAGGGCACTGGAGCTGGAGGTCTCGTCCAGCCCTCCCCATCTC  
 GGGGCTGCTGTGGACGGCGCTGCCTCAGGCACCTCCTGTGAAACCTGCCCTACTGTGTT  
 TAACCTGTTGCTCCAGGATGCAATTGATAGGAGGGGGCGGCAAGGGCTGGGCTTGTGACAATC  
 TGCCCTTCACCCACATGGCCTTGCCTCGGTGCCCTGACTGTCAAGGGAGGGCCAGGGAGGCAGAG  
 CGGGAGGGAGTCTCAGGAGGAGGCTGCCCTGAGGGGCTGGGGAGGGGTAACCTCATGAGGACCA  
 GGGTGGAGCTGAGAAGAGGAGGAGGTGGGGCTGGAGGTGCTGGTAGCTGAGGGGACGGGCAAG  
 TGAGAGGGAGGGAGGGAAAGTCCTGGGAGGATCCTGAGCTGCTGTCAGTCAACCAACTAAT  
 CAGTTCTAGATTGAGGGAGGGCAAGGCACCAACAATCAGAAATGGGGCTTTCGGGGAGGGC  
 GCCTAGTCCCCCAGCTTAAGCAGCCAGGAGGGACCTGCATCTAACATCTGGGTGCCATGG  
 CAATGGCATGCCCTCAGCTACTGTATGCCCGACCCCGCAGAGGCAGAATGAACCCATAGG

**Fig. 4 cont.**

GAGCTGATCGTAATGTTATCATGTTACTTCCCCACCCCTACATTGGAAATAAAAATAAGGA  
ATTTTATTCTCACTTCTGTGTTCTGCACGCCAATGCCAGGCCATGGTATTGGGTGATAGAT  
GAGGCCCTCTAGCTGGGCCTGGCACCAAGGAGGGTCCCCATGCTGCATCTCTGTATCCC  
CTCCCTCCCTGTGGCCTCCACCCGCCCTCCCTGCTGCCCTGTGAAATTCAATTCTGGGCC  
GGAACCTGGTGGAAATGACCCAAAACATTGCCATCTCCCTCTCAGCAGCCGACCCCAG  
CCCAATTCTAAACAGGGCTGAGAGCCACCTCTCAGCAGCTGACCCCTACCCAAGGAGGGTGGC  
ATGGAGGGGCTTGAGAGACTCTCTTAACATCCTCCCCCCCCAGCTGTCCTCCCAAGTGCAT  
CTGCCCTCCCATCCCTGGGCCAGCCAGCTCACAGAGGCCAGGCCAACAGAATTCTGGCC  
TCCTTGAAGGGGCTGGAGAAGGCCGGAGCAGTGGCTACGCCCTGTAATCCAGCACTTGGG  
AGGCTGAGGCCGGCAGATCACAAAGTCAGAGATTGAGACCATCTGGCAACATGGTGAACACC  
CCGTCTCTACTAAAAATACAAAAATTAGGCCGGGTGCGGTGGCTACGCCCTGTAATCCAGCAC  
TTTGGGAGGCCGAGGCCGGCAGATCACGAGTCAGGAGATCAAGACCCTGGTAACACGGT  
GAAACCCCGTCTACTAAAAATCACAAAAATTAGCTGGCGAGGTGGCGGGTGCCTGTAGTCC  
CAGCTACGTGGAGGCTGAGGCAAGAGAATGGCGTGAACCCCGCGGGCAGAGCCTGCAGAGA  
GCTGAGATCACACCACTGTACTCCAGCCTGGCGACAGCGAGACTCCGTCTAAAAAAAAAAA  
AAAAAAAAATTAGCTGGGCATGGTGGTGCCTGCAGTCCCAGCTACTCAGGAGGCTGAGACG  
GGAGAAATCGCTTGAACCTGGGAGGCAGAGGTTGCAGTGAAGCCAAGATCGCTCACTCCAGCCTAG  
CGACAGAGTGAGACTCCATCTCAAATAATAAATTAAATTAAATTAAATTAAATT

## Fig. 5

### FADS3 cDNA

GGCCGCGGGCGGCAGGGCGGGGCCGGAGCAGCAGGGCGGGCGGAGGCAGGGGCCGGCCCCGGAGCGCTC  
TTCGCTTCCCTCGGGTCTTGCTCGAACCTCGGCCACCGCCTGGGATCCCAGGACTCGTGCCTG  
GCAGCATGGGCGGGCGTCGGGAGCCGGACCGCGGGAGGGACCCCGCGCACGCCGGGGCACCGCT  
GCCCACTTCTGCTGGGAGCAGATCCCGCGCACGACCAGCCGGCGACAAGTGGCTGGTCATC  
GAGCGCCGCGTCTACGACATCAGCCGTGGGACAGCGGACCCAGGGGAGCCGGCTCATCG  
GCCACCACGGCGCTGAGGAGCAGCCACGGATGCCCTCCGTGCCTTCATCAAGATCTCAATTG  
GCGCAAGTTCTACAGCCCCCTGTTGAGAGCTGGCTCCGAAGAACCCAGCCAGGATGGA  
CCCCTGAATGCCAGCTGGTCAGGGACTTCCGAGGCCCTGCACCAGCAGCCGAGGACATGAAGC  
TGTTTGATGCCAGTCCCACCTCTTTGCTTCTACTGGGCCACATCCTGCCATGGAGGTGCT  
GCCCTGGCTCCTTATCTACCTCCTGGCTGGCTGGGCCAGTGCCCTGGCCCTTCATC  
CTGGCCATCTCAGGCTCAGTCTGGTGTCTGCAGCATGACCTGGGCCATGCCCTCATCTTCA  
AGAAGTCCTGGTGAACCACGTGGCCAGAAAGTTCGTATGGGCAGCTAAAGGGCTTCCGC  
CCACTGGTGAACTTCCGCCACTTCAGCACCAGCCAAGCCAAACATCTTCCACAAAGACCCA  
GACGTGACGGTGGCGCCCGTCTTCCCTCTGGGGAGTCATCCGTCGAGTATGGCAAGAAC  
GCAGATAACCTACCCCTACAAACCAGCAGCACCTGTACTCTTCCGTATGGCCCGCCTGCTCAC  
CCTGGTGAACTTGAAGTGGAAAATCTGGCGTACATGCTGGTGTGCATGCAGTGGCCGATTG  
CTCTGGGCCAGCTCTATGCCGCTTCTTCTACCTCCCTTCTACGGCGTCCCTG  
GGGTGCTGCTCTTCTTGCTGTCAGGGCTGGAAAGCCACTGGTCTGGATCACACA  
GATGAACCACATCCCCAAGGAGATGGCCACGAGAACGACCCGGACTGGTCAGCTCAGCTG  
GCAGCCACCTGCAACGTGGAGCCCTACTTTCACCAACTGGTCAAGCGGGCACCTCAACTTCC  
AGATCGAGCACCCACCTCTCCCCAGGATGCCAGACACAACACTACAGCCGGTGGCCCCGCTGGT  
CAAGTCGCTGTGTCAGACGGCCCTCAGCTACGAAGTGAAGCCTTCCCTACCGCGCTGGT  
GACATCGTCAGGTCCCTGAAGAAGTCTGGTACATCTGGCTGGACGCCAACCTCCATCAGTGA  
GGCAACACCCAGGCGGGCAGAGAAGGGCTCAGGGCACCAAGCCAGCCCCGGGGAT  
CGATACCCCCACCCTCCACTGGCCAGCCTGGGGTGCCCTGCCCTCCTGGTACTGTTG  
CTTCCCTGGCCCCCTCACATGTGATTAGCAGCCCTATGGCCTGGCTGGCCTGATGG  
GACAGGGTAGAGGGAAGGTGAGCATAGCACATTCTAGAGCGAGAATTGGGGAAAGCTGT  
TATTTTATATAAAATACATTAGATGT

**Fig. 6****FADS1**

Met	Ala	Pro	Asp	Pro	Val	Ala	Ala	Glu	Thr	Ala	Ala	Gln	Gly	Pro	Thr
1					5				10					15	
Pro	Arg	Tyr	Phe	Thr	Trp	Asp	Glu	Val	Ala	Gln	Arg	Ser	Gly	Cys	Glu
				20				25					30		
Glu	Arg	Trp	Leu	Val	Ile	Asp	Arg	Lys	Val	Tyr	Asn	Ile	Ser	Glu	Phe
				35				40				45			
Thr	Arg	Arg	His	Pro	Gly	Gly	Ser	Arg	Val	Ile	Ser	His	Tyr	Ala	Gly
				50				55			60				
Gln	Asp	Ala	Thr	Asp	Pro	Phe	Val	Ala	Phe	His	Ile	Asn	Lys	Gly	Leu
				65			70			75			80		
Val	Lys	Lys	Tyr	Met	Asn	Ser	Leu	Leu	Ile	Gly	Glu	Leu	Ser	Pro	Glu
				85				90				95			
Gln	Pro	Ser	Phe	Glu	Pro	Thr	Lys	Asn	Lys	Glu	Leu	Thr	Asp	Glu	Phe
				100			105				110				
Arg	Glu	Leu	Arg	Ala	Thr	Val	Glu	Arg	Met	Gly	Leu	Met	Lys	Ala	Asn
				115			120				125				
His	Val	Phe	Phe	Leu	Leu	Tyr	Leu	Leu	His	Ile	Leu	Leu	Leu	Asp	Gly
				130			135				140				
Ala	Ala	Trp	Leu	Thr	Leu	Trp	Val	Phe	Gly	Thr	Ser	Phe	Leu	Pro	Phe
				145			150			155			160		
Leu	Leu	Cys	Ala	Val	Leu	Leu	Ser	Ala	Val	Gln	Ala	Gln	Ala	Gly	Trp
				165			170				175				
Leu	Gln	His	Asp	Phe	Gly	His	Leu	Ser	Val	Phe	Ser	Thr	Ser	Lys	Trp
				180			185				190				
Asn	His	Leu	Leu	His	His	Phe	Val	Ile	Gly	His	Leu	Lys	Gly	Ala	Pro
				195			200				205				
Ala	Ser	Trp	Trp	Asn	His	Met	His	Phe	Gln	His	His	Ala	Lys	Pro	Asn
				210			215				220				
Cys	Phe	Arg	Lys	Asp	Pro	Asp	Ile	Asn	Met	His	Pro	Phe	Phe	Phe	Ala
				225			230			235			240		
Leu	Gly	Lys	Ile	Leu	Ser	Val	Glu	Leu	Gly	Lys	Gln	Lys	Lys	Tyr	
				245			250				255				
Met	Pro	Tyr	Asn	His	Gln	His	Lys	Tyr	Phe	Phe	Leu	Ile	Gly	Pro	Pro
				260			265				270				
Ala	Leu	Leu	Pro	Leu	Tyr	Phe	Gln	Trp	Tyr	Ile	Phe	Tyr	Phe	Val	Ile
				275			280				285				

**Fig. 6 cont.**

Gln Arg Lys Lys Trp Val Asp Leu Ala Trp Met Ile Thr Phe Tyr Val  
290 295 300

Arg Phe Phe Leu Thr Tyr Val Pro Leu Leu Gly Leu Lys Ala Phe Leu  
305 310 315 320

Gly Leu Phe Phe Ile Val Arg Phe Leu Glu Ser Asn Trp Phe Val Trp  
325 330 335

Val Thr Gln Met Asn His Ile Pro Met His Ile Asp His Asp Arg Asn  
340 345 350

Met Asp Trp Val Ser Thr Gln Leu Gln Ala Thr Cys Asn Val His Lys  
355 360 365

Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu  
370 375 380

His His Leu Phe Pro Thr Met Pro Arg His Asn Tyr His Lys Val Ala  
385 390 395 400

Pro Leu Val Gln Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Ser  
405 410 415

Lys Pro Leu Leu Ser Ala Phe Ala Asp Ile Ile His Ser Leu Lys Glu  
420 425 430

Ser Gly Gln Leu Trp Leu Asp Ala Tyr Leu His Gln  
435 440

**Fig. 7**

FADS2

Met	Gly	Lys	Gly	Gly	Asn	Gln	Gly	Glu	Gly	Ala	Ala	Glu	Arg	Glu	Vai
1					5					10				15	
Ser	Val	Pro	Thr	Phe	Ser	Trp	Glu	Glu	Ile	Gln	Lys	His	Asn	Leu	Arg
						20				25				30	
Thr	Asp	Arg	Trp	Leu	Val	Ile	Asp	Arg	Lys	Val	Tyr	Asn	Ile	Thr	Lys
						35			40				45		
Trp	Ser	Ile	Gln	His	Pro	Gly	Gly	Gln	Arg	Val	Ile	Gly	His	Tyr	Ala
						50			55				60		
Gly	Glu	Asp	Ala	Thr	Asp	Ala	Phe	Arg	Ala	Phe	His	Pro	Asp	Leu	Glu
	65					70				75				80	
Phe	Val	Gly	Lys	Phe	Leu	Lys	Pro	Leu	Leu	Ile	Gly	Glu	Leu	Ala	Pro
	85							90						95	
Glu	Glu	Pro	Ser	Gln	Asp	His	Gly	Lys	Asn	Ser	Lys	Ile	Thr	Glu	Asp
	100							105						110	
Phe	Arg	Ala	Leu	Arg	Lys	Thr	Ala	Glu	Asp	Met	Asn	Leu	Phe	Lys	Thr
	115						120						125		
Asn	His	Val	Phe	Phe	Leu	Leu	Leu	Leu	Ala	His	Ile	Ile	Ala	Leu	Glu
	130						135						140		
Ser	Ile	Ala	Trp	Phe	Thr	Val	Phe	Tyr	Phe	Gly	Asn	Gly	Trp	Ile	Pro
	145					150				155				160	
Thr	Leu	Ile	Thr	Ala	Phe	Val	Leu	Ala	Thr	Ser	Gln	Ala	Gln	Ala	Gly
		165							170				175		
Trp	Leu	Gln	His	Asp	Tyr	Gly	His	Leu	Ser	Val	Tyr	Arg	Lys	Pro	Lys
		180						185					190		
Trp	Asn	His	Leu	Val	His	Lys	Phe	Val	Ile	Gly	His	Leu	Lys	Gly	Ala
		195						200					205		
Ser	Ala	Asn	Trp	Trp	Asn	His	Arg	His	Phe	Gln	His	His	Ala	Lys	Pro
		210					215						220		
Asn	Ile	Phe	His	Lys	Asp	Pro	Asp	Val	Asn	Met	Leu	His	Val	Phe	Val
	225					230				235				240	
Leu	Gly	Glu	Trp	Gln	Pro	Ile	Glu	Tyr	Gly	Lys	Lys	Lys	Leu	Lys	Tyr
			245						250					255	
Leu	Pro	Tyr	Asn	His	Gln	His	Glu	Tyr	Phe	Phe	Leu	Ile	Gly	Pro	Pro
			260					265					270		
Leu	Leu	Ile	Pro	Met	Tyr	Phe	Gln	Tyr	Gln	Ile	Ile	Met	Thr	Met	Ile
			275						280				285		

**Fig. 7 cont.**

Val His Lys Asn Trp Val Asp Leu Ala Trp Ala Val Ser Tyr Tyr Ile  
 290 295 300  
 Arg Phe Phe Ile Thr Tyr Ile Pro Phe Tyr Gly Ile Leu Gly Ala Leu  
 305 310 315 320  
 Leu Phe Leu Asn Phe Ile Arg Phe Leu Glu Ser His Trp Phe Val Trp  
 325 330 335  
 Val Thr Gln Met Asn His Ile Val Met Glu Ile Asp Gln Glu Ala Tyr  
 340 345 350  
 Arg Asp Trp Phe Ser Ser Gln Leu Thr Ala Thr Cys Asn Val Glu Gln  
 355 360 365  
 Ser Phe Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu  
 370 375 380  
 His His Leu Phe Pro Thr Met Pro Arg His Asn Leu His Lys Ile Ala  
 385 390 395 400  
 Pro Leu Val Lys Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Glu  
 405 410 415  
 Lys Pro Leu Leu Arg Ala Leu Leu Asp Ile Ile Arg Ser Leu Lys Lys  
 420 425 430  
 Ser Gly Lys Leu Trp Leu Asp Ala Tyr Leu His Lys  
 435 440

**Fig. 8****FADS3**

Met Gly Gly Val Gly Glu Pro Gly Pro Arg Glu Gly Pro Ala Gln Pro  
 1                       5                       10                       15  
  
 Gly Ala Pro Leu Pro Thr Phe Cys Trp Glu Gln Ile Arg Ala His Asp  
 20                      25                       30  
  
 Gln Pro Gly Asp Lys Trp Leu Val Ile Glu Arg Arg Val Tyr Asp Ile  
 35                      40                       45  
  
 Ser Arg Trp Ala Gln Arg His Pro Gly Gly Ser Arg Leu Ile Gly His  
 50                      55                       60  
  
 His Gly Ala Glu Asp Ala Thr Asp Ala Phe Arg Ala Phe His Gln Asp  
 65                      70                       75                       80  
  
 Leu Asn Phe Val Arg Lys Phe Leu Gln Pro Leu Leu Ile Gly Glu Leu  
 85                      90                       95  
  
 Ala Pro Glu Glu Pro Ser Gln Asp Gly Pro Leu Asn Ala Gln Leu Val  
 100                     105                       110  
  
 Glu Asp Phe Arg Ala Leu His Gln Ala Ala Glu Asp Met Lys Leu Phe  
 115                     120                       125  
  
 Asp Ala Ser Pro Thr Phe Phe Ala Phe Leu Leu Gly His Ile Leu Ala  
 130                     135                       140  
  
 Met Glu Val Leu Ala Trp Leu Leu Ile Tyr Leu Leu Gly Pro Gly Trp  
 145                     150                       155                       160  
  
 Val Pro Ser Ala Leu Ala Ala Phe Ile Leu Ala Ile Ser Gln Ala Gln  
 165                     170                       175  
  
 Ser Trp Cys Leu Gln His Asp Leu Gly His Ala Ser Ile Phe Lys Lys  
 180                     185                       190  
  
 Ser Trp Trp Asn His Val Ala Gln Lys Phe Val Met Gly Gln Leu Lys  
 195                     200                       205  
  
 Gly Phe Ser Ala His Trp Trp Asn Phe Arg His Phe Gln His His Ala  
 210                     215                       220  
  
 Lys Pro Asn Ile Phe His Lys Asp Pro Asp Val Thr Val Ala Pro Val  
 225                     230                       235                       240  
  
 Phe Leu Leu Gly Glu Ser Ser Val Glu Tyr Gly Lys Lys Lys Arg Arg  
 245                     250                       255  
  
 Tyr Leu Pro Tyr Asn Gln Gln His Leu Tyr Phe Phe Leu Ile Gly Pro  
 260                     265                       270  
  
 Pro Leu Leu Thr Leu Val Asn Phe Glu Val Glu Asn Leu Ala Tyr Met  
 275                     280                       285

**Fig. 8 cont.**

Leu Val Cys Met Gln Trp Ala Asp Leu Leu Trp Ala Ala Ser Phe Tyr  
290 295 300

Ala Arg Phe Phe Leu Ser Tyr Leu Pro Phe Tyr Gly Val Pro Gly Val  
305 310 315 320

Leu Leu Phe Phe Val Ala Val Arg Val Leu Glu Ser His Trp Phe Val  
325 330 335

Trp Ile Thr Gln Met Asn His Ile Pro Lys Glu Ile Gly His Glu lys  
340 345 350

His Arg Asp Trp Val Ser Ser Gln Leu Ala Ala Thr Cys Asn Val Glu  
355 360 365

Pro Ser Leu Phe Thr Asn Trp Phe Ser Gly His Leu Asn Phe Gln Ile  
370 375 380

Glu His His Leu Phe Pro Arg Met Pro Arg His Asn Tyr Ser Arg Val  
385 390 395 400

Ala Pro Leu Val Lys Ser Leu Cys Ala Lys His Gly Leu Ser Tyr Glu  
405 410 415

Val Lys Pro Phe Leu Thr Ala Leu Val Asp Ile Val Arg Ser Leu lys  
420 425 430

Lys Ser Gly Asp Ile Trp Leu Asp Ala Tyr Leu His Gln  
435 440 445

## Fig. 9

### Oligonucleotide primers to amplify FADS1 cDNA

TU12-R5 (5'-CGCCTGACAGCCCCTGCT-3')

TU12-F10 (5'-CAGGTGGCCAATCACAAAAT-3')

TU12-R7 (5'-CTCAAAGTGGAACCATCTGCTA-3')

TU12-F9 (5'-GGAAACCCAGTCCATGTTCC-3')

TU12-R6 (5'-CCTGGGCCTTTCTTCATAGT-3')

TU12-F5 (5'-CTCAAGCTCCCCTCTGCCT-3')

### Oligonucleotide primers to amplify FADS2 cDNA

TU13-R4 (5'-TCAGAACATAACCTGCGC-3')

TU13-F7 (5'-CCAGTTCACCAATCAGCAGG-3')

TU13-R3 (5'-CCCTGCTGATTGGTGAAC-3')

TU13-F4 (5'-TGTAGGGCAGGTATTCAGC-3')

TU13-R2 (5'-AGCCCATCGAGTACGGCAA-3')

TU13-F1 (5'-CCTCAGAACAAAAGCCCATC-3')

### Oligonucleotide primers to amplify FADS3 cDNA

TU19-R2 (5'-TCTTGCTCGGACCTCGGC-3')

TU19-F2 (5'-GTGATCCACACGAACCAGTG-3')

TU19-R3 (5'-GAAGAACCCAGCCAGGATG-3')

TU19-F3 (5'-ACAGCTTCCCCAATTCTC-3')



European Patent  
Office

### PARTIAL EUROPEAN SEARCH REPORT

Application Number

which under Rule 45 of the European Patent Convention EP 99 10 4664  
shall be considered, for the purposes of subsequent  
proceedings, as the European search report

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	"AC A1394672" EMBL DATABASE, 5 February 1999 (1999-02-05), XP002111712 Heidelberg * the whole document * ---	1-12	C12N15/53 C12N15/11 C12N15/85 C12N9/02 C12N5/10 C12Q1/02
X	WO 98 46763 A (TURMOND JENNIFER ; CALGENE LLC (US); ABBOTT LAB (US); KNUTZON DEBO) 22 October 1998 (1998-10-22) * see esp. SEQ ID NOS: 27-40 ---	1-12, 17, 18	C07K16/40 A61K39/395 A61K38/44 A01K67/027 G01N33/50 G01N33/53
X	CHO H P ET AL: "Cloning, expression, and nutritional regulation of the mammalian Delta-6 desaturase." JOURNAL OF BIOLOGICAL CHEMISTRY, (1999 JAN 1) 274 (1) 471-7., XP002111713 * the whole document * ---	1-12, 17, 18	
X	"AC 060426" EMBL DATABASE, 1 August 1998 (1998-08-01), XP002111714 Heidelberg * the whole document * ---	7-10 -/-	
INCOMPLETE SEARCH			TECHNICAL FIELDS SEARCHED (Int.Cl.7)
			C12N C12Q C07K A61K A01K G01N
<p>The Search Division considers that the present application, or one or more of its claims, does/do not comply with the EPC to such an extent that a meaningful search into the state of the art cannot be carried out, or can only be carried out partially, for these claims.</p> <p>Claims searched completely:</p> <p>Claims searched incompletely</p> <p>Claims not searched:</p> <p>Reason for the limitation of the search: see sheet C</p>			
Place of search	Date of completion of the search	Examiner:	
THE HAGUE	10 August 1999	Kania, T	
CATEGORY OF CITED DOCUMENTS			
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document			
T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

European Patent  
OfficeINCOMPLETE SEARCH  
SHEET CApplication Number  
EP 99 10 4664

Although claim 17 is directed to a method of treatment of the human/animal body (Article 52(4) EPC), the search has been carried out and based on the alleged effects of the compound/composition.

-----  
Claim(s) searched incompletely:  
16

Reason for the limitation of the search:

Claims 14 and 15 were only interpreted and searched with reference to the use of the present molecules and vectors in these assays.  
Claim 16 could not be searched completely due to the lack of characterization of the claimed subject matter.



European Patent  
Office

Application Number  
EP 99 10 4664

### CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing more than ten claims.

- Only part of the claims have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims and for those claims for which claims fees have been paid, namely claim(s):
  
- No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for the first ten claims.

### LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

see sheet B

- All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.
  
- As all searchable claims could be searched without effort justifying an additional fee, the Search Division did not invite payment of any additional fee.
  
- Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid, namely claims:
  
- None of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims, namely claims:



The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

1. Claims: 1-18 partially

An isolated polynucleotide selected from the group consisting of polynucleotides having at least 65%, preferably 80% homology with a polynucleotide encoding a polypeptide of SEQ ID NO:4, comprising variants, under stringent conditions hybridizing molecules, complementary molecules, and oligonucleotides comprising at least 15 consecutive nucleotides of said sequence, preferably the polynucleotide of SEQ ID NO:1. Vectors, host cells, and transgenic organisms comprising said sequences.

A polypeptide comprising a sequence having at least 65%, more preferably 85% homology to SEQ ID NO:4, variants thereof, and a peptide comprising at least 15 consecutive amino acids thereof. A process for producing said polypeptide using said host cells and DNA sequences. Antibodies against said polypeptides, and their use in diagnosis and therapy.

An oligonucleotide primer having a sequence selected from the group of nucleotide sequences of SEQ ID NOs:7-12. A method of screening for modulators in known assays using constructs or of screening for interacting proteins or factors using state of the art technologies, as well as a method of screening chemical libraries comprising transformed cell lines, both methods employing the said sequences, vectors, or host cells.

A compound which alters or reacts with at least one epitope of the proteins and which is obtained by said methods. Pharmaceutical compositions comprising as an effective component an effective amount of said peptides.

2. Claims: 1-18 partially

idem for SEQ ID NOs:2,5,13-18

3. Claims: 1-18 partially

idem for SEQ ID NOs:3,6,19-22



European Patent  
Office

## PARTIAL EUROPEAN SEARCH REPORT

Application Number

EP 99 10 4664

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
			TECHNICAL FIELDS SEARCHED (Int.Cl.7)
X	"AC 060427" EMBL DATABASE, 1 August 1998 (1998-08-01), XP002111715 Heidelberg * the whole document * ---	7-10	
A	OLGA SAYANOVA ET AL: "Expression of a borage desaturase cDNA containing an N-terminal cytochrome b5 domain results in the accumulation of high levels of Delta6-desaturated fatty acids in transgenic tobacco" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 94, no. 94, 15 April 1997 (1997-04-15), pages 4211-4216 XP002106758 ISSN: 0027-8424 * the whole document *	1-18	
A	MITCHELL, ANDREW G. ET AL: "A novel cytochrome b-5-like domain is linked to the carboxyl terminus of the <i>Saccharomyces cerevisiae</i> DELTA-9 fatty acid desaturase." JOURNAL OF BIOLOGICAL CHEMISTRY, (1995) VOL. 270, NO. 50, PP. 29766-29772 , XP002111716 * the whole document *	1-18	
A	WO 96 02561 A (GEN HOSPITAL CORP ;GENETICS INST (US)) 1 February 1996 (1996-02-01) * the whole document *	14,16	
A	WO 99 04262 A (MYELOS NEUROSCIENCES CORP) 28 January 1999 (1999-01-28) * see esp. claims *	15,16	

**ANNEX TO THE EUROPEAN SEARCH REPORT  
ON EUROPEAN PATENT APPLICATION NO.**

EP 99 10 4664

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.  
The members are as contained in the European Patent Office EDP file on  
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

10-08-1999

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO 9846763 A	22-10-1998	AU	6961698 A	11-11-1998
		AU	7114798 A	11-11-1998
		WO	9846764 A	22-10-1998
-----	-----	-----	-----	-----
WO 9602561 A	01-02-1996	EP	0773952 A	21-05-1997
		JP	10504713 T	12-05-1998
-----	-----	-----	-----	-----
WO 9904262 A	28-01-1999	AU	8480798 A	10-02-1999
-----	-----	-----	-----	-----

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82